

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
19 April 2001 (19.04.2001)

PCT

(10) International Publication Number
WO 01/27857 A2

(51) International Patent Classification?: **G06F 19/00**

(21) International Application Number: **PCT/US00/28413**

(22) International Filing Date: 13 October 2000 (13.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/159,176 13 October 1999 (13.10.1999) US
60/217,658 10 July 2000 (10.07.2000) US
60/217,251 10 July 2000 (10.07.2000) US
09/663,968 19 September 2000 (19.09.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/159,176 (CIP)
Filed on	13 October 1999 (13.10.1999)
US	60/217,658 (CIP)
Filed on	10 July 2000 (10.07.2000)
US	09/663,968 (CIP)
Filed on	19 September 2000 (19.09.2000)
US	60/217,251 (CIP)
Filed on	10 July 2000 (10.07.2000)

(71) Applicant (for all designated States except US): **SEQUENOM, INC. [US/US]**; 11555 Sorrento Valley Road, San Diego, CA 92121-1331 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BRAUN, Andreas** [DE/US]; 11237-6 Carmel Creek Road, San Diego, CA 92130 (US). **KÖSTER, Hubert** [DE/US]; 8636-C Via Mallorca Drive, La Jolla, CA 92037 (US). **VAN**

DEN BOOM, Dirk [DE/DE]; Eppendorfer Weg 205 D, D-20253 Hamburg (DE). **PING, Yip** [US/US]; 3641 Copley Avenue, San Diego, CA 92116 (US). **RODI, Charlie** [US/US]; 13823 Recuerdo Drive, Del Mar, CA 92014 (US). **HE, Liyan** [CN/US]; 10948 Creek Bridge Place, San Diego, CA 92128 (US). **CHIU, Norman** [CA/US]; 1128 Caminito Alvarez, San Diego, CA 92126 (US). **JURINKE, Christian** [DE/DE]; Rombergstrasse 22, 20255 Hamburg (DE).

(74) Agents: **SEIDMAN, Stephanie, L.** et al.; Heller Ehrman White & McAuliffe, Suite 700, 4250 Executive Square, La Jolla, CA 92037 (US).

(81) Designated States (national): AE, AG, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/27857 A2

(54) Title: METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS

(57) Abstract: Process and methods for creating a database of genomic samples from healthy human donors, methods that use the database to identify and correlate polymorphic genetic markers and other markers with diseases and conditions are provided.

-1-

METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS

RELATED APPLICATIONS

Benefit of priority to the following applications is claimed herein:

- 5 U.S. provisional application Serial No. 60/217,658 to Andreas Braun, Hubert Koster; Dirk Van den Boom, filed July 10, entitled "METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS"; U.S. provisional application Serial No. 60/159,176 to Andreas Braun, Hubert Koster, Dirk Van den Boom, filed October 10, 1999, entitled "METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS"; U.S. provisional application Serial No. 60/217,251, filed July 10, 2000, to Andreas Braun, entitled "POLYMORPHIC KINASE ANCHOR PROTEIN GENE SEQUENCES, POLYMORPHIC KINASE ANCHOR PROTEINS AND METHODS OF DETECTING
- 15 POLYMORPHIC KINASE ANCHOR PROTEINS AND NUCLEIC ACIDS ENCODING THE SAME"; and U.S. application Serial No. 09/663,968, to Ping Yip, filed September 19, 2000, entitled "METHOD AND DEVICE FOR IDENTIFYING A BIOLOGICAL SAMPLE."

Where permitted that above-noted applications and provisional applications are incorporated by reference in their entirety.

FIELD OF THE INVENTION

Process and methods for creating a database of genomic samples from healthy human donors. Methods that use the database to identify and correlate with polymorphic genetic markers and other markers with diseases and conditions are provided.

BACKGROUND

Diseases in all organisms have a genetic component, whether inherited or resulting from the body's response to environmental stresses, such as viruses and toxins. The ultimate goal of ongoing genomic research is to use this information to develop new ways to identify, treat and potentially cure these diseases. The first step has been to screen disease tissue and identify genomic changes at the level of individual samples. The identification of these "disease"

-2-

markers has then fueled the development and commercialization of diagnostic tests that detect these errant genes or polymorphisms. With the increasing numbers of genetic markers, including single nucleotide polymorphisms (SNPs), microsatellites, tandem repeats, newly mapped introns and exons, the challenge 5 to the medical and pharmaceutical communities is to identify genotypes which not only identify the disease but also follow the progression of the disease and are predictive of an organism's response to treatment.

Currently the pharmaceutical and biotechnology industries find a disease and then attempt to determine the genomic basis for the disease. This approach 10 is time consuming and expensive and in many cases involves the investigator guessing as to what pathways might be involved in the disease.

Genomics

Presently the two main strategies employed in analyzing the available genomic information are the technology driven reverse genetics brute force 15 strategy and the knowledge-based pathway oriented forward genetics strategy. The brute force approach yields large databases of sequence information but little information about the medical or other uses of the sequence information. Hence this strategy yields intangible products of questionable value. The knowledge-based strategy yields small databases that contain a lot of 20 information about medical uses of particular DNA sequences and other products in the pathway and yield tangible products with a high value.

Polymorphisms

Polymorphisms have been known since 1901 with the identification of blood types. In the 1950's they were identified on the level of proteins using 25 large population genetic studies. In the 1980's and 1990's many of the known protein polymorphisms were correlated with genetic loci on genomic DNA. For example, the gene dose of the apolipoprotein E type 4 allele was correlated with the risk of Alzheimer's disease in late onset families (see, e.g., Corder *et al.* (1993) *Science* 261: 921-923; mutation in blood coagulation factor V was 30 associated with resistance to activated protein C (see, e.g., Bertina *et al.* (1994) *Nature* 369:64-67); resistance to HIV-1 infection has been shown in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene (see,

-3-

e.g., Samson *et al.* (1996) *Nature* 382:722-725); and a hypermutable tract in antigen presenting cells (APC, such as macrophages), has been identified in familial colorectal cancer in individuals of Ashkenzi jewish background (see, e.g., Laken *et al.* (1997) *Nature Genet.* 17:79-83). There may be more than three
5 million polymorphic sites in the human genome. Many have been identified, but not yet characterized or mapped or associated with a marker.

Single nucleotide polymorphisms (SNPs)

Much of the focus of genomics has been in the identification of SNPs, which are important for a variety of reasons. They allow indirect testing
10 (association of haplotypes) and direct testing (functional variants). They are the most abundant and stable genetic markers. Common diseases are best explained by common genetic alterations, and the natural variation in the human population aids in understanding disease, therapy and environmental interactions.

15 Currently, the only available method to identify SNPs in DNA is by sequencing, which is expensive, difficult and laborious. Furthermore, once a SNP is discovered it must be validated to determine if it is a real polymorphism and not a sequencing error. Also, discovered SNPs must then be evaluated to determine if they are associated with a particular phenotype. Thus, there is a
20 need to develop new paradigms for identifying the genomic basis for disease and markers thereof. Therefore, it is an object herein to provide methods for identifying the genomic basis of disease and markers thereof.

SUMMARY

25 Databases and methods using the databases are provided herein. The databases comprise sets of parameters associated with subjects in populations selected only on the basis of being healthy (*i.e.*, where the subjects are mammals, such as humans, they are selected based upon apparent health and no detectable infections). The databases can be sorted based upon one or more of the selected parameters.

30 The databases are preferably relational databases, in which an index that represents each subject serves to relate parameters, which are the data, such as age, ethnicity, sex, medical history, etc. and ultimately genotypic information,

-4-

that was inputted into and stored in the database. The database can then be sorted according to these parameters. Initially, the parameter information is obtained from a questionnaire answered by each subject from whom a body tissue or body fluid sample is obtained. As additional information about each 5 sample is obtained, this information can be entered into the database and can serve as a sorting parameter.

The databases obtained from healthy individuals have numerous uses, such as correlating known polymorphisms with a phenotype or disease. The databases can be used to identify alleles that are deleterious, that are beneficial, 10 and that are correlated with diseases.

For purposes herein, genotypic information can be obtained by any method known to those of skill in the art, but is preferably obtained using mass spectrometry.

Also provided herein, is a new use for existing databases of subjects and 15 genotypic and other parameters, such as age, ethnicity, race, and gender. Any database can be sorted according to the methods herein and alleles that exhibit statistically significant correlations with any of the sorting parameters can be identified. It is noted, however, is noted, that the databases provided herein and randomly selected databases will perform better in these methods, since disease-based databases suffer numerous limitations, including their relatively small size, 20 the homogeneity of the selected disease population, and the masking effect of the polymorphism associated with the markers for which the database was selected. Hence, the healthy database provided herein, provides advantages not heretofore recognized or exploited. However, the methods provided herein can 25 be used with a selected database, including disease-based databases, with or without sorting for the discovery and correlation of polymorphisms. In addition, the databases provided herein represent a greater genetic diversity than the unselected databases typically utilized for the discovery of polymorphisms and thus allow for the enhanced discovery and correlation of polymorphisms.

30 The databases provided herein can be used for taking an identified polymorphism, and ascertaining whether it changes in frequency when the data is sorted according to a selected parameter.

-5-

One use of these methods is correlating a selected marker with a particular parameter by following the occurrence of known genetic markers and then, having made this correlation, determining or identifying correlations with diseases. Examples of this use are p53 and Lipoprotein Lipase polymorphism.

- 5 As exemplified herein, known markers are shown to have particular correlation with certain groups, such as a particular ethnicity or race or one sex. Such correlations will then permit development of better diagnostic tests and treatment regimens.

These methods are valuable for identifying one or more genetic markers 10 whose frequency changes within the population as a function of age, ethnic group, sex or some other criteria. This can allow the identification of previously unknown polymorphisms and ultimately a gene or pathway involved in the onset and progression of disease.

The databases and methods provided herein permit, among other things, 15 identification of components, particularly key components, of a disease process by understanding its genetic underpinnings and also permit an understanding of processes, such as individual drug responses. The databases and methods provided herein also can be used in methods involving elucidation of pathological pathways, in developing new diagnostic assays, identifying new potential drug 20 targets, and in identifying new drug candidates.

The methods and databases can be used with experimental procedures, including, but are not limited to, *in silico* SNP identification, *in vitro* SNP identification/verification, genetic profiling of large populations, and in biostatistical analyses and interpretations.

25 Also provided herein, are combinations that contain a database provided herein and a biological sample from a subject in the database, and preferably biological samples from all subjects or a plurality of subjects in the database. Collections of the tissue and body fluid samples are also provided.

Also, provided herein, are methods for determining a genetic marker that 30 correlates with age, comprising identifying a polymorphism and determining the frequency of the polymorphism with increasing age in a healthy population.

-6-

Further provided herein are methods for determining whether a genetic marker correlates with susceptibility to morbidity, early mortality, or morbidity and early mortality, comprising identifying a polymorphism and determining the frequency of the polymorphism with increasing age in a healthy population.

- 5 Any of the methods herein described can be used out in a multiplex format.

Also provided are an apparatus and process for accurately identifying genetic information. It is another object of the herein that genetic information be extracted from genetic data in a highly automated manner. Therefore, to
10 overcome the deficiencies in the known conventional systems, a method and apparatus for identifying a biological sample is proposed.

Briefly, the method and system for identifying a biological sample generates a data set indicative of the composition of the biological sample. In a particular example, the data set is DNA spectrometry data received from a mass
15 spectrometer. The data set is denoised, and a baseline is deleted. Since possible compositions of the biological sample may be known, expected peak areas may be determined. Using the expected peak areas, a residual baseline is generated to further correct the data set. Probable peaks are then identifiable in the corrected data set, which are used to identify the composition of the
20 biological sample. In a disclosed example, statistical methods are employed to determine the probability that a probable peak is an actual peak, not an actual peak, or that the data too inconclusive to call.

Advantageously, the method and system for identifying a biological sample accurately makes composition calls in a highly automated manner. In
25 such a manner, complete SNP profile information, for example, may be collected efficiently. More importantly, the collected data is analyzed with highly accurate results. For example, when a particular composition is called, the result may be relied upon with great confidence. Such confidence is provided by the robust computational process employed .

30 **DESCRIPTION OF THE DRAWINGS**

-7-

Figure 1 depicts an exemplary sample bank. Panel 1 shows the samples as a function of sex and ethnicity. Panel 2 shows the caucasians as a function of age. Panel 3 shows the Hispanics as a function of age.

Figures 2A and 2C show an age- and sex-distribution of the 291S allele
5 of the lipoprotein lipase gene in which a total of 436 males and 589 females were investigated. Figure 2B shows an age distribution for the 436 males.

Figure 3 is an exemplary questionnaire for population-based sample banking.

Figure 4 depicts processing and tracking of blood sample components.

10 Figure 5 depicts the allelic frequency of "sick" alleles and "healthy" alleles as a function of age. It is noted that the relative frequency of healthy alleles increases in a population with increasing age.

Figure 6 depicts the age-dependent distribution of ApoE genotypes (see, Schächter *et al.* (1994) *Nature Genetics* 6:29-32).

15 Figure 7A-D depicts age-related and genotype frequency of the p53 (tumor suppressor) codon 72 among the caucasian population in the database. *R72 and *P72 represent the frequency of the allele in the database population. R72, R72P, and P72 represent the genotypes of the individuals in the population. The frequency of the homozygous P72 allele drops from 6.7% to 3.7% with
20 age.

Figure 8 depicts the allele and genotype frequencies of the p21 S31R allele as a function of age.

Figure 9 depicts the frequency of the FVII Allele 353Q in pooled versus individual samples.

25 Figure 10 depicts the frequency of the CETP (cholesterol ester transfer protein) allele in pooled versus individual samples

Figure 11 depicts the frequency of the plasminogen activator inhibitor-1 (PAI-1) 5G in pooled versus individual samples

30 Figure 12 shows mass spectra of the samples and the ethnic diversity of the PAI-1 alleles.

Figure 13 shows mass spectra of the samples and the ethnic diversity of the CETP 405 alleles.

-8-

Figure 14 shows mass spectra of the samples and the ethnic diversity of the Factor VII 353 alleles.

Figure 15 shows ethnic diversity of PAI-1, CETP and Factor VII using the pooled DNA samples.

5 Figure 16 shows the p53-Rb pathway and the relationships among the various factors in the pathway.

Figure 17, which is a block diagram of a computer constructed to provide and process the databases described herein, depicts a typical computer system for storing and sorting the databases provided herein and practicing the methods
10 provided herein.

Figure 18 is a flow diagram that illustrates the processing steps performed using the computer illustrated in Figure 17, to maintain and provide access to the databases for identifying polymorphic genetic markers.

Figure 19 is a histogram showing the allele and genotype distribution in
15 the age and sex stratified Caucasian population for the AKAP10-1 locus. Bright green bars show frequencies in individuals younger than 40 years. Dark green bars show frequencies in individuals older than 60 years.

Figure 20 is a histogram showing the allele and genotype distribution in the age and sex stratified Caucasian population for the AKAP10-5 locus. Bright
20 green bars show frequencies in individuals younger than 40 years; dark green bars show frequencies in individuals older than 60 years.

Figure 21 is a histogram showing the allele and genotype distribution in the age and sex stratified Caucasian population for the h-msrA locus. Genotype difference between male age groups is significant. Bright green bars show
25 frequencies in individuals younger than 40 years. Dark green bars show frequencies in individuals older than 60 years.

Figure 22A-D is a sample data collection questionnaire used for the healthy database.

Figure 23 is a flowchart showing processing performed by the computing
30 device of Figure 24 when performing genotyping of sense strands and antisense strands from assay fragments.

-9-

Figure 24 is a block diagram showing a system in accordance with the present invention;

Figure 25 is a flowchart of a method of identifying a biological sample in accordance with the present invention;

5 Figure 26 is a graphical representation of data from a mass spectrometer;

Figure 27 is a diagram of wavelet transformation of mass spectrometry data;

Figure 28 is a graphical representation of wavelet stage 0 hi data;

Figure 29 is a graphical representation of stage 0 noise profile;

10 Figure 30 is a graphical representation of generating stage noise standard deviations;

Figure 31 is a graphical representation of applying a threshold to data stages;

Figure 32 is a graphical representation of a sparse data set;

15 Figure 33 is a formula for signal shifting;

Figure 34 is a graphical representation of a wavelet transformation of a denoised and shifted signal;

Figure 35 is a graphical representation of a denoised and shifted signal;

Figure 36 is a graphical representation of removing peak sections;

20 Figure 37 is a graphical representation of generating a peak free signal ;

Figure 38 is a block diagram of a method of generating a baseline correction;

Figure 39 is a graphical representation of a baseline and signal;

Figure 40 is a graphical representation of a signal with baseline removed;

25 Figure 41 is a table showing compressed data;

Figure 42 is a flowchart of method for compressing data;

Figure 43 is a graphical representation of mass shifting;

Figure 44 is a graphical representation of determining peak width;

Figure 45 is a graphical representation of removing peaks;

30 Figure 46 is a graphical representation of a signal with peaks removed;

Figure 47 is a graphical representation of a residual baseline;

-10-

Figure 48 is a graphical representation of a signal with residual baseline removed;

Figure 49 is a graphical representation of determining peak height;

Figure 50 is a graphical representation of determining signal-to-noise for
5 each peak;

Figure 51 is a graphical representation of determining a residual error for each peak;

Figure 52 is a graphical representation of peak probabilities;

Figure 53 is a graphical representation of applying an allelic ratio to peak
10 probability;

Figure 54 is a graphical representation of determining peak probability

Figure 55 is a graphical representation of calling a genotype;

Figure 56 is a flowchart showing a statistical procedure for calling a genotype;

15 Figure 57 is a flowchart showing processing performed by the computing device of Figure 1 when performing standardless genotyping; and

Figure 58 is graphical representation of applying an allelic ratio to peak probability for standardless genotype processing.

DETAILED DESCRIPTION

20 Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of ordinary skill in the art to which this invention belongs. All patents, applications, published applications and other publications and sequences from GenBank and other
25 databases referred to herein throughout the disclosure are incorporated by reference in their entirety.

As used herein, a biopolymer includes, but is not limited to, nucleic acid, proteins, polysaccharides, lipids and other macromolecules. Nucleic acids include DNA, RNA, and fragments thereof. Nucleic acids may be derived from
30 genomic DNA, RNA, mitochondrial nucleic acid, chloroplast nucleic acid and other organelles with separate genetic material.

-11-

As used herein, morbidity refers to conditions, such as diseases or disorders, that compromise the health and well-being of an organism, such as an animal. Morbidity susceptibility or morbidity-associated genes are genes that, when altered, for example, by a variation in nucleotide sequence, facilitate the expression of a specific disease clinical phenotype. Thus, morbidity susceptibility genes have the potential, upon alteration, of increasing the likelihood or general risk that an organism will develop a specific disease.

5

- As used herein, mortality refers to the statistical likelihood that an organism, particularly an animal, will not survive a full predicted lifespan.
- 10 Hence, a trait or a marker, such as a polymorphism, associated with increased mortality is observed at a lower frequency in older than younger segments of a population.

As used herein, a polymorphism, e.g. genetic variation, refers to a variation in the sequence of a gene in the genome amongst a population, such as allelic variations and other variations that arise or are observed. Thus, a polymorphism refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. These differences can occur in coding and non-coding portions of the genome, and can be manifested or detected as differences in nucleic acid sequences, gene expression, including,

15

20 for example transcription, processing, translation, transport, protein processing, trafficking, DNA synthesis, expressed proteins, other gene products or products of biochemical pathways or in post-translational modifications and any other differences manifested amongst members of a population. A single nucleotide polymorphism (SNP) refers to a polymorphism that arises as the result of a single base change, such as an insertion, deletion or change in a base.

25

A polymorphic marker or site is the locus at which divergence occurs. Such site may be as small as one base pair (an SNP). Polymorphic markers include, but are not limited to, restriction fragment length polymorphisms, variable number of tandem repeats (VNTR's), hypervariable regions,

30 minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats and other repeating patterns, simple sequence repeats and insertional elements, such as Alu. Polymorphic forms also are manifested as different mendelian

-12-

alleles for a gene. Polymorphisms may be observed by differences in proteins, protein modifications, RNA expression modification, DNA and RNA methylation, regulatory factors that alter gene expression and DNA replication, and any other manifestation of alterations in genomic nucleic acid or organelle nucleic acids.

5 As used herein, a healthy population, refers to a population of organisms, including but are not limited to, animals, bacteria, viruses, parasites, plants, eubacteria, and others, that are disease free. The concept of disease-free is a function of the selected organism. For example, for mammals it refers to a subject not manifesting any disease state. Practically a healthy subject, when
10 human, is defined as human donor who passes blood bank criteria to donate blood for eventual use in the general population. These criteria are as follows: free of detectable viral, bacterial, mycoplasma, and parasitic infections; not anemic; and then further selected based upon a questionnaire regarding history (see Figure 3). Thus, a healthy population represents an unbiased population of
15 sufficient health to donate blood according to blood bank criteria, and not further selected for any disease state. Typically such individuals are not taking any medications. For plants, for example, it is a plant population that does not manifest diseases pathology associated with plants. For bacteria it is a bacterial population replicating without environmental stress, such as selective agents,
20 heat and other pathogens.

As used herein, a healthy database (or healthy patient database) refers to a database of profiles of subjects that have not been pre-selected for any particular disease. Hence, the subjects that serve as the source of data for the database are selected, according to predetermined criteria, to be healthy. In
25 contrast to other such databases that have been pre-selected for subjects with a particular disease or other characteristic, the subjects for the database provided herein are not so-selected. Also, if the subjects do manifest a disease or other condition, any polymorphism discovered or characterized should be related to an independent disease or condition. In a preferred embodiment, where the
30 subjects are human, a healthy subject manifests no disease symptoms and meets criteria, such as those set by blood banks for blood donors.

-13-

Thus, the subjects for the database are a population of any organism, including, but are not limited to, animals, plants, bacteria, viruses, parasites and any other organism or entity that has nucleic acid. Among preferred subjects are mammals, preferably, although not necessarily, humans. Such a database can
5 capture the diversity of the a population, thus providing for discovery of rare polymorphisms.

As used herein, a profile refers to information relating to, but not limited to and not necessarily including all of, age, sex, ethnicity, disease history, family history, phenotypic characteristics, such as height and weight and other relevant
10 parameters. A sample collect information form is shown in Figure 22, which illustrates profile intent.

As used herein, a disease state is a condition or abnormality or disorder that may be inherited or result from environmental stresses, such as toxins, bacterial, fungal and viral infections.

15 As used herein, set of non-selected subjects means that the subjects have not been pre-selected to share a common disease or other characteristic. They can be selected to be healthy as defined herein.

As used herein, a phenotype refers to a set of parameters that includes any distinguishable trait of an organism. A phenotype can be physical traits and
20 can be, in instances in which the subject is an animal, a mental trait, such as emotional traits. Some phenotypes can be determined by observation elicited by questionnaires (see, e.g., Figures 3 and 22) or by referring to prior medical and other records. For purposes herein, a phenotype is a parameter around which the database can be sorted.

25 As used herein, a parameter is any input data that will serve as a basis for sorting the database. These parameters will include phenotypic traits, medical histories, family histories and any other such information elicited from a subject or observed about the subject. A parameter may describe the subject, some historical or current environmental or social influence experienced by the
30 subject, or a condition or environmental influence on someone related to the subject. Paramaters include, but are not limited to, any of those described herein, and known to those of skill in the art.

-14-

As used herein, haplotype refers refers to two or polymorphism located on a single DNA strand. Hence, haplotyping refers to identification of two or more polymorphisms on a single DNA strand. Haplotypes can be indicative of a phenotype. For some disorders a single polymorphism may suffice to indicate a trait; for others a plurality (i.e., a haplotype) may be needed. Haplotyping can be performed by isolating nucleic acid and separating the strands. In addition, when using enzymes such as certain nucleases, that produce, different size fragments from each strand, strand separation is not needed for haplotyping.

As used herein, used herein, pattern with reference to a mass spectrum or mass spectrometric analyses, refers to a characteristic distribution and number of signals (such peaks or digital representations thereof).

As used herein, signal in the context of a mass spectrum and analysis thereof refers to the output data, which the number or relative number of molecules having a particular mass. Signals include "peaks" and digital representations thereof.

As used herein, adaptor, when used with reference to haplotyping use of ligase, refers to a nucleic acid that specifically hybridizes to a polymorphism of interest. An adaptor can be partially double-stranded. An adaptor complex is formed when an adaptor hybridizes to its target.

As used herein, a target nucleic acid refers to any nucleic acid of interest in a sample. It can contain one or more nucleotides.

As used herein, standardless analysis refers to a determination based upon an internal standard. For example, the frequency of a polymorphism can be determined herein by comparing signals within a single mass spectrum.

As used herein, amplifying refers to means for increasing the amount of a polymer, especially nucleic acids. Based on the 5' and 3' primers that are chosen, amplification also serves to restrict and define the region of the genome which is subject to analysis. Amplification can be by any means known to those skilled in the art, including use of the polymerase chain reaction (PCR) etc.

Amplification, e.g., PCR must be done quantitatively when the frequency of polymorphism is required to be determined.

-15-

As used herein, cleaving refers to non-specific and specific fragmentation of a biopolymer.

As used herein, multiplexing refers to the simultaneous detection of more than one polymorphism. Methods for performing multiplexed reactions,

- 5 particularly in conjunction with mass spectrometry are known (see, e.g., U.S. Patent Nos. 6,043,031, 5,547,835 and International PCT application No. WO 97/37041).

As used herein, reference to mass spectrometry encompasses any suitable mass spectrometric format known to those of skill in the art. Such formats

- 10 include, but are not limited to, Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI (see, e.g., published International PCT application No.99/57318 and U.S. Patent No. 5,118,937), Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof. MALDI, particular UV and IR, are among the preferred formats.

- 15 As used herein, mass spectrum refers to the presentation of data obtained from analyzing a biopolymer or fragment thereof by mass spectrometry either graphically or encoded numerically.

As used herein, a blood component is a component that is separated from blood and includes, but is not limited to red blood cells and platelets, blood

- 20 clotting factors, plasma, enzymes, plasminogen, immunoglobulins. A cellular blood component is a component of blood, such as a red blood cell, that is a cell. A blood protein is a protein that is normally found in blood. Examples of such proteins are blood factors VII and VIII. Such proteins and components are well-known to those of skill in the art.

- 25 As used herein, plasma can be prepared by any method known to those of skill in the art. For example, it can be prepared by centrifuging blood at a force that pellets the red cells and forms an interface between the red cells and the buffy coat, which contains leukocytes, above which is the plasma. For example, typical platelet concentrates contain at least about 10% plasma.

- 30 Blood may be separated into its components, including, but not limited to, plasma, platelets and red blood cells by any method known to those of skill in the art. For example, blood can be centrifuged for a sufficient time and at a

sufficient acceleration to form a pellet containing the red blood cells. Leukocytes collect primarily at the interface of the pellet and supernatant in the buffy coat region. The supernatant, which contains plasma, platelets, and other blood components, may then be removed and centrifuged at a higher acceleration,

- 5 whereby the platelets pellet.

As used herein, p53 is a cell cycle control protein that assesses DNA damage and acts as a transcription factor regulation gene which control cell growth, DNA repair and apoptosis. The p53 mutations have been found in a wide variety of different cancers, including all of the different types of leukemia, 10 with varying frequency. The loss of normal p53 functions results in genomic instability and uncontrolled growth of the host cell.

As used herein, p21 is a cyclin-dependent kinase inhibitor, associated with G1 phase arrest of normal cells. Expression triggers apoptosis or programmed cell death and has been associated with Wilms' tumor, a pediatric 15 kidney cancer.

As used herein, Factor VII is a serine protease involved the extrinsic blood coagulation cascade. This factor is activated by thrombin and works with tissue factor (Factor III) in the processing of Factor X to Factor Xa. Evidence has supported an association between polymorphisms in the gene and increase 20 Factor VII activity which can result in an elevated risk of ischemic cardiovascular disease including myocardial infarction.

As used herein, a relational database stores information in a form representative of matrices, such as two-dimensional tables, including rows and columns of data, or higher dimensional matrices. For example, in one 25 embodiment, the relational database has separate tables each with a parameter. The tables are linked with a record number, which also acts as an index. The database can be searched or sorted by using data in the tables and is stored in any suitable storage medium, such as floppy disk, CD rom disk, hard drive or other suitable medium.

30 As used herein, a bar codes refers any array of optically readable marks of any desired size and shape that are arranged in a reference context or frame of, preferably, although not necessarily, one or more columns and one or more

-17-

rows. For purposes herein, the bar code refers to any symbology, not necessary "bar" but may include dots, characters or any symbol or symbols.

As used herein, symbology refers to an identifier code or symbol, such as a bar code, that is linked to a sample. The index will reference each such

- 5 symbology. The symbology is any code known or designed by the user. The symbols are associated with information stored in the database. For example, each sample can be uniquely identified with an encoded symbology. The parameters, such as the answers to the questions and subsequent genotypic and other information obtained upon analysis of the samples is included in the
- 10 database and associated with the symbology. The database is stored on any suitable recording medium, such as a hard drive, a floppy disk, a tape, a CD ROM, a DVD disk and any other suitable medium.

DATABASES

Human genotyping is currently dependent on collaborations with

- 15 hospitals, tissues banks and research institutions that provide samples of disease tissue. This approach is based on the concept that the onset and/or progression of diseases can be correlated with the presence of a polymorphisms or other genetic markers. This approach does not consider that disease correlated with the presence of specific markers and the absence of specific markers. It is
- 20 shown herein that identification and scoring of the appearance and disappearance of markers is possible only if these markers are measured in the background of healthy subjects where the onset of disease does not mask the change in polymorphism occurrence. Databases of information from disease populations suffer from small sample size, selection bias and heterogeneity. The
- 25 databases provided herein from healthy populations solve these problems by permitting large sample bands, simple selection methods and diluted heterogeneity.

- 30 Provided herein are first databases of parameters, associated with non-selected, particularly healthy, subjects. Also provided are combinations of the databases with indexed samples obtained from each of the subjects. Further provided are databases produced from the first databases. These contain in addition to the original parameters information, such as genotypic information,

-18-

including, but are not limited to, genomic sequence information, derived from the samples.

The databases, which are herein designated healthy databases, are so-designated because they are not obtained from subjects pre-selected for a 5 particular disease. Hence, although individual members may have a disease, the collection of individuals is not selected to have a particular disease.

The subjects from whom the parameters are obtained comprise either a set of subjects who are randomly selected across, preferably, all populations, or are pre-selected to be disease-free or healthy. As a result, the database is not 10 selected to be representative of any pre-selected phenotype, genotype, disease or other characteristic. Typically the number of subjects from which the database is prepared is selected to produce statistically significant results when used in the methods provided herein. Preferably, the number of subjects will be greater than 100, more preferably greater than 200, yet more preferably greater 15 than 1000. The precise number can be empirically determined based upon the frequency of the parameter(s) that be used to sort the database. Generally the population can have at least 50, at least 100, at least 200, at least 500, at least 1000, at least 5000 or at least 10,000 or more subjects.

Upon identification of a collection of subjects, information about each 20 subject is recorded and associated with each subject as a database. The information associated with each of the subjects, includes, but is not limited to, information related to historical characteristics of the subjects, phenotypic characteristics and also genotypic characteristics, medical characteristics and any other traits and characteristics about the subject that can be determined. 25 This information will serve as the basis for sorting the database.

In an exemplary embodiment, the subjects are mammals, such as humans, and the information relates to one or more of parameters, such as age, sex, medical history, ethnicity and any other factor. Such information, when the animals are humans, for example, can be obtained by a questionnaire, and by 30 observations about the individual, such as hair color, eye color and other characteristics. Genotypic information will be obtained from tissue or other body and body fluid samples from the subject.

-19-

The healthy genomic database can include profiles and polymorphisms from healthy individuals from a library of blood samples where each sample in the library is an individual and separate blood or other tissue sample. Each sample in the database is profiled as to the sex, age, ethnic group, and disease history of the donor.

5 The databases are generated by first identifying healthy populations of subjects and obtaining information about each subject that will serve as the sorting parameters for the database. This information is preferably entered into a storage medium, such as the memory of a computer.

10 The information obtained about each subject in a population used for generating the database is stored in a computer memory or other suitable storage medium. The information is linked to an identifier associated with each subject. Hence the database will identify a subject, for example by a datapoint representative of a bar code, and then all information, such as the information 15 from a questionnaire, regarding the individual is associated with the datapoint. As the information is collected the database is generated.

Thus, for example, profile information, such as subject histories obtained from questionnaires, is collected in the database. The resulting database can be sorted as desired, using standard software, such as by age, sex and/or ethnicity.

20 An exemplary questionnaire for subjects from whom samples are to be obtained is shown in Figures 22A-D. Each questionnaire preferably is identified by a bar code, particularly a machine readable bar code for entry into the database. After a subject provides data and is deemed to be healthy (*i.e.*, meets standards for blood donation), the data in the questionnaire is entered into the database and is 25 associated with the bar code. A tissue, cell or blood sample is obtained from the subject.

Figure 4 exemplifies processing and tracking of blood sample components. Each component is tracked with a bar code, dated, is entered into the database and associated with the subject and the profile of the subject.

30 Typically, the whole blood is centrifuged to produce plasma, red blood cells (which pellet) and leukocytes found in the buffy coat which layers in between.

-20-

Various samples are obtained and coded with a bar code and stored for use as needed.

Samples are collected from the subjects. The samples include, but are not limited to, tissues, cells, and fluids, such as nucleic acid, blood, plasma,

- 5 amniotic fluid, synovial fluid, urine, saliva, aqueous humor, sweat, sperm samples and cerebral spinal fluid. It is understood that the particular set of samples depends upon the organisms in the population.

- Once samples are obtained the collection can be stored and, in preferred embodiments, each sample is indexed with an identifier, particularly a machine 10 readable code, such as a bar code. For analyses, the samples or components of the samples, particularly biopolymers and small molecules, such as nucleic acids and/or proteins and metabolites, are isolated.

- After samples are analyzed, this information is entered into the database in the memory of the storage medium and associated with each subject. This 15 information includes, but is not limited to, genotypic information. Particularly, nucleic acid sequence information and other information indicative of polymorphisms, such as masses of PCR fragments, peptide fragment sequences or masses, spectra of biopolymers and small molecules and other indicia of the structure or function of a gene, gene product or other marker from which the 20 existence of a polymorphism within the population can be inferred.

- In an exemplary embodiment, a database can be derived from a collection of blood samples. For example, Figure 1 (see, also Figure 10) shows the status of a collection of over 5000 individual samples. The samples were processed in the laboratory following SOP (standard operating procedure) guidelines. Any 25 standard blood processing protocol may be used.

For the exemplary database described herein, the following criteria were used to select subjects:

No testing is done for infectious agents.

Age: At least 17 years old

- 30 Weight: Minimum of 110 pounds

Permanently Disqualified:

History of hepatitis (after age 11)

-21-

Leukemia Lymphoma

Human immunodeficiency virus (HIV),AIDS

Chronic kidney disease

Temporarily Disqualified:

- 5 Pregnancy - until six weeks after delivery, miscarriage or abortion
- Major surgery or transfusions - for one year
- Mononucleosis - until complete recovery
- Prior whole blood donation - for eight weeks
- Antibiotics by injection for one week; by mouth, for forty-eight hours,
- 10 except antibiotics for skin complexion;

5 year Deferment:

Internal cancer and skin cancer if it has been removed, is healed and there is no recurrence

- These correspond to blood bank criteria for donating blood and represent a
- 15 healthy population as defined herein for a human healthy database.

Structure of the database

- Any suitable database structure and format known to those of skill in the art may be employed. For example, a relational database is a preferred format in which data is stored as matrices or tables of the parameters linked by an indexer
- 20 that identifies each subject. Software for preparing and manipulating, including sorting the database, can be readily developed or adapted from commercially available software, such as Microsoft Access.

Quality control

- Quality control procedures can be implemented. For example, after
- 25 collection of samples, the quality of the collection in the bank can be assessed. For example, mix-up of samples can be checked by testing for known markers, such as sex. After samples are separated by ethnicity, samples are randomly tested for a marker associated with a particular ethnicity, such as HLA DQA1 group specific component, to assess whether the samples have been properly
 - 30 sorted by ethnic group. An exemplary sample bank is depicted in Figure 4.

-22-

Obtaining genotypic data and other parameters for the database

After informational and historical parameters are entered into the database, material from samples obtained from each subject, is analyzed.

Analyzed material include proteins, metabolites, nucleic acids, lipids and any

- 5 other desired constituent of the material. For example, nucleic acids, such as genomic DNA, can be analyzed by sequencing.

Sequencing can be performed using any method known to those of skill in the art. For example, if a polymorphism is identified or known, and it is desired to assess its frequency or presence among the subjects in the database, the

- 10 region of interest from each sample can be isolated, such as by PCR or restriction fragments, hybridization or other suitable method known to those of skill in the art and sequenced. For purposes herein, sequencing analysis is preferably effected using mass spectrometry (see, e.g., U.S. Patent Nos. 5,547,835, 5,622,824, 5,851,765, and 5,928,906). Nucleic acids can also be
15 sequence by hybridization (see, e.g., U.S. Patent Nos. 5,503,980, 5,631,134, 5,795,714) and including analysis by mass spectrometry (see, U.S. application Serial Nos. 08/419,994 and 09/395,409).

In other detection methods, it is necessary to first amplify prior to identifying the allelic variant. Amplification can be performed, e.g., by PCR

- 20 and/or LCR, according to methods known in the art. In one embodiment, genomic DNA of a cell is exposed to two PCR primers and amplification for a number of cycles sufficient to produce the required amount of amplified DNA. In preferred embodiments, the primers are located between 150 and 350 base pairs apart.

- 25 Alternative amplification methods include: self sustained sequence replication (Guatelli, J. C. et al., 1990, Proc. Natl. Acad. Sci. U.S.A. 87:1874-1878), transcriptional amplification system (Kwoh, D. Y. et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:1173-1177), Q-Beta Replicase (Lizardi, P. M. et al., 1988, Bio/Technology 6:1197), or any other nucleic acid amplification
30 method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially

-23-

useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

Nucleic acids can also be analyzed by detection methods and protocols, particularly those that rely on mass spectrometry (see, e.g., U.S. Patent No.

- 5 5,605,798, 6,043,031, allowed copending U.S. application Serial No. 08/744,481, U.S. application Serial No. 08/990,851 and International PCT application No. WO 99/31273, International PCT application No. WO 98/20019). These methods can be automated (see, e.g., copending U.S. application Serial No. 09/285,481 and published International PCT application No.
- 10 10 PCT/US00/08111, which describes an automated process line). Preferred among the methods of analysis herein are those involving the primer oligo base extension (PROBE) reaction with mass spectrometry for detection (described herein and elsewhere, see e.g., U.S. Patent No. 6,043,031; see, also U.S. application Serial Nos. 09/287,681, 09/287,682, 09/287,141 and 09/287,679,
- 15 15 allowed copending U.S. application Serial No. 08/744,481, International PCT application No. PCT/US97/20444, published as International PCT application No. WO 98/20019, and based upon U.S. application Serial Nos. 08/744,481, 08/744,590, 08/746,036, 08/746,055, 08/786,988, 08/787,639, 08/933,792, 08/746,055, 08/786,988 and 08/787,639; see, also U.S. application Serial No.
- 20 20 09/074,936, U.S. Patent No. 6,024,925, and U.S. application Serial Nos. 08/746,055 and 08/786,988, and published International PCT application No. WO 98/20020)

A preferred format for performing the analyses is a chip based format in which the biopolymer is linked to a solid support, such as a silicon or silicon-

- 25 25 coated substrate, preferably in the form of an array. More preferably, when analyses are performed using mass spectrometry, particularly MALDI, small nanoliter volumes of sample are loaded on, such that the resulting spot is about, or smaller than, the size of the laser spot. It has been found that when this is achieved, the results from the mass spectrometric analysis are quantitative. The
- 30 30 area under the signals in the resulting mass spectra are proportional to concentration (when normalized and corrected for background). Methods for preparing and using such chips are described in U.S. Patent No. 6,024,925, co-

-24-

pending U.S. application Serial Nos. 08/786,988, 09/364,774, 09/371,150 and 09/297,575; see, also U.S. application Serial No. PCT/US97/20195, which published as WO 98/20020. Chips and kits for performing these analyses are commercially available from SEQUENOM under the trademark MassARRAY.

5 MassArray relies on the fidelity of the enzymatic primer extension reactions combined with the miniaturized array and MALDI-TOF (Matrix-Assisted Laser Desorption Ionization-Time of Flight) mass spectrometry to deliver results rapidly. It accurately distinguishes single base changes in the size of DNA fragments associated with genetic variants without tags.

10 The methods provided herein permit quantitative determination of alleles. The areas under the signals in the mass spectra can be used for quantitative determinations. The frequency is determined from the ratio of the signal to the total area of all of the spectrum and corrected for background. This is possible because of the PROBE technology as described in the above applications
15 incorporated by reference herein.

Additional methods of analyzing nucleic acids include amplification-based methods including polymerase chain reaction (PCR), ligase chain reaction (LCR), mini-PCR, rolling circle amplification, autocatalytic methods, such as those using Q β replicase, TAS, 3SR, and any other suitable method known to those of skill
20 in the art.

Other methods for analysis and identification and detection of polymorphisms, include but are not limited to, allele specific probes, Southern analyses, and other such analyses.

The methods described below provide ways to fragment given amplified
25 or non-amplified nucleotide sequences thereby producing a set of mass signals when mass spectrometry is used to analyze the fragment mixtures. Amplified fragments are yielded by standard polymerase chain methods (US 4,683,195 and 4,683,202). The fragmentation method involves the use of enzymes that cleave single or double strands of DNA and enzymes that ligate
30 DNA. The cleavage enzymes can be glycosylases, nickases, and site-specific and non site-specific nucleases with the most preferred enzymes being glycosylases, nickases, and site-specific nucleases.

Glycosylase Fragmentation Method

DNA glycosylases specifically remove a certain type of nucleobase from a given DNA fragment. These enzymes can thereby produce abasic sites, which can be recognized either by another cleavage enzyme, cleaving the exposed

- 5 phosphate backbone specifically at the abasic site and producing a set of nucleobase specific fragments indicative of the sequence, or by chemical means, such as alkaline solutions and or heat. The use of one combination of a DNA glycosylase and its targeted nucleotide would be sufficient to generate a base specific signature pattern of any given target region.
- 10 Numerous DNA glycosylases are known. For example, a DNA glycosylase can be uracil-DNA glycosylase (UDG), 3-methyladenine DNA glycosylase, 3-methyladenine DNA glycosylase II, pyrimidine hydrate-DNA glycosylase, FaPy-DNA glycosylase, thymine mismatch-DNA glycosylase, hypoxanthine-DNA glycosylase, 5-Hydroxymethyluracil DNA glycosylase (HmUDG), 5-
- 15 Hydroxymethylcytosine DNA glycosylase, or 1,N6-ethenoadenine DNA glycosylase (see, e.g., U.S. Patent Nos. 5,536,649, 5,888, 795, 5,952,176 and 6,099,553, International PCT application Nos. WO 97/03210, WO 99/54501; see, also, Eftedal et al. (1993) Nucleic Acids Res 21:2095-2101, Bjelland and Seeberg (1987) Nucleic Acids Res. 15:2787-2801, Saparbaev et al.
- 20 (1995) Nucleic Acids Res. 23:3750-3755, Bessho (1999) Nucleic Acids Res. 27:979-983) corresponding to the enzyme's modified nucleotide or nucleotide analog target. A preferred glycosylase is uracil-DNA glycosylase (UDG).

- Uracil, for example, can be incorporated into an amplified DNA molecule by amplifying the DNA in the presence of normal DNA precursor nucleotides (e.g. dCTP, dATP, and dGTP) and dUTP. When the amplified product is treated with UDG, uracil residues are cleaved. Subsequent chemical treatment of the products from the UDG reaction results in the cleavage of the phosphate backbone and the generation of nucleobase specific fragments. Moreover, the separation of the complementary strands of the amplified product prior to 30 glycosylase treatment allows complementary patterns of fragmentation to be generated. Thus, the use of dUTP and Uracil DNA glycosylase allows the generation of T specific fragments for the complementary strands, thus providing

-26-

information on the T as well as the A positions within a given sequence. Similar to this, a C-specific reaction on both (complementary) strands (i.e. with a C-specific glycosylase) yields information on C as well as G positions within a given sequence if the fragmentation patterns of both amplification strands are 5 analyzed separately. Thus, with the glycosylase method and mass spectrometry, a full series of A, C, G and T specific fragmentation patterns can be analyzed.

Nickase Fragmentation Method

A DNA nickase, or DNase, can be used recognize and cleave one strand 10 of a DNA duplex. Numerous nickases are known. Among these, for example, are nickase NY2A nickase and NYS1 nickase (Megabase) with the following cleavage sites:

NY2A: 5'...R AG...3'

3'...Y TC...5' where R = A or G and Y = C or T

15 NYS1: 5'... CC[A/G/T]...3'

3'... GG[T/C/A]...5'.

Fen-Ligase Fragmentation Method

The Fen-ligase method involves two enzymes: Fen-1 enzyme and a ligase. The Fen-1 enzyme is a site-specific nuclease known as a "flap" endonuclease 20 (US 5,843,669, 5,874,283, and 6,090,606). This enzymes recognizes and cleaves DNA "flaps" created by the overlap of two oligonucleotides hybridized to a target DNA strand. This cleavage is highly specific and can recognize single base pair mutations, permitting detection of a single homologue from an individual heterozygous at one SNP of interest and then genotyping that 25 homologue at other SNPs occurring within the fragment. Fen-1 enzymes can be Fen-1 like nucleases e.g. human, murine, and Xenopus XPG enzymes and yeast RAD2 nucleases or Fen-1 endonucleases from, for example, *M. jannaschii*, *P. furiosus*, and *P. woesei*. Among preferred enzymes are the Fen-1 enzymes.

The ligase enzyme forms a phosphodiester bond between two double 30 stranded nucleic acid fragments. The ligase can be DNA Ligase I or DNA Ligase III (see, e.g., U.S. Patent Nos. US 5,506,137, 5,700,672, 5,858,705 and 5,976,806; see, also, Waga, et al. (1994) J. Biol. Chem. 269:10923-10934, Li

et al. (1994) Nucleic Acids Res. 22:632-638, Arrand et al. (1986) J. Biol. Chem. 261:9079-9082, Lehman (1974) Science 186:790-797, Higgins and Cozzarelli (1979) Methods Enzymol. 68:50-71, Lasko et al. (1990) Mutation Res. 236:277-287, and Lindahl and Barnes (1992) Ann. Rev. Biochem. 61:251-281).

- 5 Thermostable ligase (Epicenter Technologies), where "thermostable" denotes that the ligase retains activity even after exposure to temperatures necessary to separate two strands of DNA, are among preferred ligases for use herein.

Type IIS Enzyme Fragmentation Method

- 10 Restriction enzymes bind specifically to and cleave double-stranded DNA at specific sites within or adjacent to a particular recognition sequence. These enzymes have been classified into three groups (e.g. Types I, II, and III) as known to those of skill in the art. Because of the properties of type I and type III enzymes, they have not been widely used in molecular biological applications.
- 15 Thus, for this invention type II enzymes are preferred. Of the thousands of restriction enzymes known in the arts, there are 179 different type II specificities. Of the 179 unique type II restriction endonucleases, 31 have a 4-base recognition sequence, 11 have a 5-base recognition sequence, 127 have a 6-base recognition sequence, and 10 have recognition sequences of greater than 20 six bases (US 5,604,098). Of category type II enzymes, type IIS is preferred.

Type IIS enzymes can be *A/lw XI*, *Bbv I*, *Bce 83*, *Bpm I*, *Bsg I*, *Bsm AI*, *Bsm FI*, *Bsa I*, *Bcc I*, *Bcg I*, *Ear I*, *Eco 57I*, *Esp 3I*, *Fau I*, *Fok I*, *Gsu I*, *Hga I*, *Mme I*, *Mbo II*, *Sap I*, and the like. The preferred type IIS enzyme is Fok I.

- The Fok I enzyme endonuclease is an exemplary well characterized member of the Type IIS class (see, e.g., U.S. Patent Nos. 5,714,330, 5,604,098, 5,436,150, 6,054,276 and 5,871,911; see, also, Szybalski et al. (1991) Gene 100:13-26, Wilson and Murray (1991) Ann. Rev. Genet. 25:585-627, Sugisaki et al. (1981) Gene 16:73-78, Podhajska and Szalski (1985) Gene 40:175-182. Fok I recognizes the sequence 5'GGATG-3' and cleaves DNA accordingly. Type IIS restriction sites can be introduced into DNA targets by incorporating the site into primers used to amplify such targets. Fragments produced by digestion with Fok I are site specific and can be analyzed by mass

-28-

spectrometry methods such as MALDI-TOF mass spectrometry, ESI-TOF mass spectrometry, and any other type of mass spectrometry well known to those of skill in the art.

Once a polymorphism has been found to correlate with a parameter
5 such as age. The possibility of false results due to allelic dropout is examined by doing comparative PCR in an adjacent region of the genome.

Analyses

In using the database, allelic frequencies can be determined across the population by analyzing each sample in the population individually, determining
10 the presence or absence of allele or marker of interest in each individual sample, and then determining the frequency of the marker in the population. The database can then be sorted (stratified) to identify any correlations between the allele and a selected parameter using standard statistical analysis. If a correlation is observed, such as a decrease in a particular marker with age or
15 correlation with sex or other parameter, then the marker is a candidate for further study, such as genetic mapping to identify a gene or pathway in which it is involved. The marker may then be correlated, for example, with a disease. Haplotyping can also be carried out. Genetic mapping can be effected using standard methods and may also require use of databases of others, such as
20 databases previously determined to be associated with a disorder.

Exemplary analyses have been performed and these are shown in the figures, and discussed herein.

Sample pooling

It has been found that using the databases provided herein, or any other
25 database of such information, substantially the same frequencies that were obtained by examining each sample separately can be obtained by pooling samples, such as in batches of 10, 20, 50, 100, 200, 500, 1000 or any other number. A precise number may be determined empirically if necessary, and can be as low as 3.

In one embodiment, the frequency of genotypic and other markers can be obtained by pooling samples. To do this a target population and a genetic variation to be assessed is selected, a plurality of samples of biopolymers are obtained from members of the population, and the biopolymer from which the 5 marker or genotype can be inferred is determined or detected. A comparison of samples tested in pools and individually and the sorted results therefrom are shown in Figure 9, which shows frequency of the factor VII Allele 353Q. Figure 10 depicts the frequency of the CETP Allele CETP in pooled versus individual samples. Figure 15 shows ethnic diversity among various ethnic groups in the 10 database using pooled DNA samples to obtain the data. Figures 12-14 show mass spectra for these samples.

Pooling of test samples has application not only to the healthy databases provided herein, but also to use in gathering data for entry into any database of subjects and genotypic information, including typical databases derived from 15 diseased populations. What is demonstrated herein, is the finding that the results achieved are statistically the same as the results that would be achieved if each sample is analyzed separately. Analysis of pooled samples by a method, such as the mass spectrometric methods provided herein, permits resolution of such data and quantitation of the results.

20 For factor VII the R53Q acid polymorphism was assessed. In Figure 9, the "individual" data represent allelic frequency observed in 92 individuals reactions. The pooled data represent the allelic frequency of the same 92 individuals pooled into a single probe reaction. The concentration of DNA in the samples of individual donors is 250 nanograms. The total concentration of DNA 25 in the pooled samples is also 250 nanograms, where the concentration of any individual DNA is 2.7 nanograms.

It also was shown that it is possible to reduce the DNA concentration of individuals in a pooled samples from 2.7 nanograms to 0.27 nanograms without any change in the quality of the spectrum or the ability to quantitate the amount 30 of sample detected. Hence low concentrations of sample may be used in the pooling methods.

-30-

Use of the databases and markers identified thereby

The successful use of genomics requires a scientific hypothesis (*i.e.*, common genetic variation, such as a SNP), a study design (*i.e.*, complex disorders), samples and technology, such as the chip-based mass spectrometric analyses (see, *e.g.*, U.S. Patent No. 5,605,798, U.S. Patent No. 5,777,324, U.S. Patent No. 6,043,031, allowed copending U.S. application Serial No. 08/744,481, U.S. application Serial No. 08/990,851, International PCT application No. WO 98/20019, copending U.S. application Serial No. 09/285,481, which describes an automated process line for analyses; see, also, U.S. application Serial Nos. 08/617,256, 09/287,681, 09/287,682, 09/287,141 and 09/287,679, allowed copending U.S. application Serial No. 08/744,481, International PCT application No. PCT/US97/20444, published as International PCT application No. WO 98/20019, and based upon U.S. application Serial Nos. 08/744,481, 08/744,590, 08/746,036, 08/746,055, 08/786,988, 08/787,639, 08/933,792, 08/746,055, 09/266,409, 08/786,988 and 08/787,639; see, also U.S. application Serial No. 09/074,936). All of these aspects can be used in conjunction with the databases provided herein and samples in the collection.

The databases and markers identified thereby can be used, for example, for identification of previously unidentified or unknown genetic markers and to identify new uses for known markers. As markers are identified, these may be entered into the database to use as sorting parameters from which additional correlations may be determined.

Previously unidentified or unknown genetic markers

The samples in the healthy databases can be used to identify new polymorphisms and genetic markers, using any mapping, sequencing, amplification and other methodologies, and in looking for polymorphisms among the population in the database. The thus-identified polymorphism can then be entered into the database for each sample, and the database sorted (stratified) using that polymorphism as a sorting parameter to identify any patterns and correlations that emerge, such as age correlated changes in the frequency of the identified marker. If a correlation is identified, the locus of the marker can be mapped and its function or effect assessed or deduced.

-31-

Thus, the databases here provide means for:

identification of significantly different allelic frequencies of genetic factors by comparing the occurrence or disappearance of the markers with increasing age in population and then associating the markers with a disease or a

5 biochemical pathway;

identification of significantly different allelic frequencies of disease

causing genetic factors by comparing the male with the female population or comparing other selected stratified populations and associating the markers with a disease or a biochemical pathway;

10 identification of significantly different allelic frequencies of disease

causing genetic factors by comparing different ethnic groups and associating the markers with a disease or a biochemical pathway that is known to occur in high frequency in the ethnic group;

profiling potentially functional variants of genes through the general

15 panmixed population stratified according to age, sex, and ethnic origin and thereby demonstrating the contribution of the variant genes to the physical condition of the investigated population;

identification of functionally relevant gene variants by gene disequilibrium analysis performed within the general panmixed population stratified according

20 to age, sex, and ethnic origin and thereby demonstrating their contribution to the physical condition of investigated population;

identification of potentially functional variants of chromosomes or parts of chromosomes by linkage disequilibrium analysis performed within the general panmixed population stratified according to age, sex, and ethnic origin and

25 thereby demonstrating their contribution to the physical condition of investigated population.

Uses of the identified markers and known markers

The databases may also be used in conjunction with known markers and sorted to identify any correlations. For example, the databases can be used for:

30 determination and evaluation of the penetrance of medically relevant polymorphic markers;

-32-

- determination and evaluation of the diagnostic specificity of medically relevant genetic factors;
- determination and evaluation of the positive predictive value of medically relevant genetic factors;
- 5 determination and evaluation of the onset of complex diseases, such as, but are not limited to, diabetes, hypertension, autoimmune diseases, arteriosclerosis, cancer and other diseases within the general population with respect to their causative genetic factors;
- delineation of the appropriate strategies for preventive disease treatment;
- 10 delineation of appropriate timelines for primary disease intervention;
- validation of medically relevant genetic factors identified in isolated populations regarding their general applicability;
- validation of disease pathways including all potential target structures identified in isolated populations regarding their general applicability; and
- 15 validation of appropriate drug targets identified in isolated populations regarding their general applicability.
- Among the diseases and disorders for which polymorphisms may be linked include, those linked to inborn errors of metabolism, acquired metabolic disorders, intermediary metabolism, oncogenesis pathways, blood clotting pathways, and DNA synthetic and repair pathways DNA repair/replication/transcription factors and activities, e.g., such as genes related to oncogenesis, aging and genes involved in blood clotting and the related biochemical pathways that are related to thrombosis, embolism, stroke, myocardial infarction, angiogenesis and oncogenesis.
- 20 25 For example, a number of diseases are caused by or involve deficient or defective enzymes in intermediary metabolism (see, e.g., Tables 1 and 2, below) that result, upon ingestion of the enzyme substrates, in accumulation of harmful metabolites that damage organs and tissues, particularly an infant's developing brain and other organs, resulting in mental retardation and other developmental disorders.

-33-

Identification of markers and genes for such disorders is of great interest.

Model systems

Several gene systems, p21, p53 and Lipoprotein Lipase polymorphism (N291S), were selected. The p53 gene is a tumor suppressor gene that is 5 mutated in diverse tumor types. One common allelic variant occurs at codon 72. A polymorphism that has been identified in the p53 gene, i.e., the R72P allele, results in an amino acid exchange, arginine to proline, at codon 72 of the gene.

Using diseased populations, it has been shown that there are ethnic differences in the allelic distribution of these alleles among African-Americans 10 and Caucasians in the U.S. The results here support this finding and also demonstrate that the results obtained with a healthy database are meaningful (see, Figure 7B).

The 291S allele leads to reduced levels of high density lipoprotein cholesterol (HDL-C) that is associated with an increased risk of males for 15 arteriosclerosis and in particular myocardial infarction (see, Reymer *et al.* (1995) *Nature Genetics* 10:28-34).

Both genetic polymorphisms were profiled within a part of the Caucasian population-based sample bank. For the polymorphism located in the lipoprotein lipase gene a total of 1025 unselected individuals (436 males and 589 females) 20 were tested. Genomic DNA was isolated from blood samples obtained from the individuals.

As shown in the Examples and figures, an exemplary database containing about 5000 subjects, answers to the questionnaire (see Figure 3), and genotypic information has been stratified. A particular known allele has been selected, and 25 the samples tested for the marker using mass spectrometric analyses, particularly PROBE (see the EXAMPLES) to identify polymorphisms in each sample. The population in the database has been sorted according to various parameters and correlations have been observed. For example, FIGURES 2A-C, show sorting of the data by age and sex for the Lipoprotein Lipase gene in the 30 Caucasian population in the database. The results show a decrease in the frequency of the allele with age in males but no such decrease in females. Other

-34-

alleles that have been tested against the database, include, alleles of p53, p21 and factor VII. Results when sorted by age are shown in the figures.

These examples demonstrate an effect of altered frequency of disease causing genetic factors within the general population. The scientific interpretation of those results allows prediction of medical relevance of polymorphic genetic alterations. In addition, conclusions can be drawn with regard to their penetrance, diagnostic specificity, positive predictive value, onset of disease, most appropriate onset of preventive strategies, and the general applicability of genetic alterations identified in isolated populations to panmixed populations.

Therefore, an age- and sex-stratified population-based sample bank that is ethnically homogenous is a suitable tool for rapid identification and validation of genetic factors regarding their potential medical utility.

**15 Exemplary computer system for creating, storing and processing the databases
Systems**

Systems, including computers, containing the databases are provided herein. The computers and databases can be used in conjunction, for example, with the APL system (see, copending U.S. application Serial No. 09/285,481),
20 which is an automated system for analyzing biopolymers, particularly nucleic acids. Results from the APL system can be entered into the database.

Any suitable computer system may be used. The computer system may be integrated into systems for sample analysis, such as the automated process line described herein (see, e.g., copending U.S. application Serial No.
25 09/285,481).

Figure 17 is a block diagram of a computer constructed in to provide and process the databases described herein. The processing that maintains the database and performs the methods and procedures may be performed on multiple computers all having a similar construction, or may be performed by a
30 single, integrated computer. For example, the computer through which data is added to the database may be separate from the computer through which the database is sorted, or may be integrated with it. In either arrangement, the

-35-

computers performing the processing may have a construction as illustrated in Figure 17.

- Figure 17 is a block diagram of an exemplary computer 1700 that maintains the database described above and performs the methods and procedures. Each computer 1700 operates under control of a central processor unit (CPU) 1702, such as a "Pentium" microprocessor and associated integrated circuit chips, available from Intel Corporation of Santa Clara, California, USA. A computer user can input commands and data from a keyboard and display mouse 1704 and can view inputs and computer output at a display 1706. The display is typically a video monitor or flat panel display device. The computer 1700 also includes a direct access storage device (DASD) 1707, such as a fixed hard disk drive. The memory 1708 typically comprises volatile semiconductor random access memory (RAM). Each computer preferably includes a program product reader 1710 that accepts a program product storage device 1712, from which the program product reader can read data (and to which it can optionally write data). The program product reader can comprise, for example, a disk drive, and the program product storage device can comprise removable storage media such as a magnetic floppy disk, an optical CD-ROM disc, a CD-R disc, a CD-RW disc, or a DVD data disc. If desired, the computers can be connected so they can communicate with each other, and with other connected computers, over a network 1713. Each computer 1700 can communicate with the other connected computers over the network 1713 through a network interface 1714 that enables communication over a connection 1716 between the network and the computer.
- The computer 1700 operates under control of programming steps that are temporarily stored in the memory 1708 in accordance with conventional computer construction. When the programming steps are executed by the CPU 1702, the pertinent system components perform their respective functions. Thus, the programming steps implement the functionality of the system as described above. The programming steps can be received from the DASD 1707, through the program product reader 1712, or through the network connection 1716. The storage drive 1710 can receive a program product, read

-36-

- programming steps recorded thereon and transfer the programming steps into the memory 1708 for execution by the CPU 1702. As noted above, the program product storage device 1710 can comprise any one of multiple removable media having recorded computer-readable instructions, including
- 5 magnetic floppy disks and CD-ROM storage discs. Other suitable program product storage devices can include magnetic tape and semiconductor memory chips. In this way, the processing steps necessary for operation can be embodied on a program product.

Alternatively, the program steps can be received into the operating

10 memory 1708 over the network 1713. In the network method, the computer receives data including program steps into the memory 1708 through the network interface 1714 after network communication has been established over the network connection 1716 by well-known methods that will be understood by those skilled in the art without further explanation. The program steps are then

15 executed by the CPU 1702 to implement the processing of the Garment Database system.

It should be understood that all of the computers of the system preferably have a construction similar to that shown in Figure 17, so that details described with respect to the Figure 17 computer 1700 will be understood to apply to all

20 computers of the system 1700. This is indicated by multiple computers 1700 shown connected to the network 1713. Any one of the computers 1700 can have an alternative construction, so long as they can communicate with the other computers and support the functionality described herein.

Figure 18 is a flow diagram that illustrates the processing steps

25 performed using the computer illustrated in Figure 17, to maintain and provide access to the databases, such as for identifying polymorphic genetic markers. In particular, the information contained in the database is stored in computers having a construction similar to that illustrated in Figure 17. The first step for maintaining the database, as indicated in Figure 18, is to identify healthy

30 members of a population. As noted above, the population members are subjects that are selected only on the basis of being healthy, and where the subjects are mammals, such as humans, they are preferably selected based upon apparent

health and the absence of detectable infections. The step of identifying is represented by the flow diagram box numbered 1802.

The next step, represented by the flow diagram box numbered 1804, is to obtain identifying and historical information and data relating to the identified members of the population. The information and data comprise parameters for each of the population members, such as member age, ethnicity, sex, medical history, and ultimately genotypic information. Initially, the parameter information is obtained from a questionnaire answered by each member, from whom a body tissue or body fluid sample also is obtained. The step of entering and storing these parameters into the database of the computer is represented by the flow diagram box numbered 1806. As additional information about each population member and corresponding sample is obtained, this information can be inputted into the database and can serve as a sorting parameter.

In the next step, represented by the flow diagram box numbered 1808, the parameters of the members are associated with an indexer. This step may be executed as part of the database storage operation, such as when a new data record is stored according to the relational database structure and is automatically linked with other records according to that structure. The step 1806 also may be executed as part of a conventional data sorting or retrieval process, in which the database entries are searched according to an input search or indexing key value to determine attributes of the data. For example, such search and sort techniques may be used to follow the occurrence of known genetic markers and then determine if there is a correlation with diseases for which they have been implicated. Examples of this use are for assessing the frequencies of the p53 and Lipoprotein Lipase polymorphisms.

Such searching of the database also may be valuable for identifying one or more genetic markers whose frequency changes within the population as a function of age, ethnic group, sex, or some other criteria. This can allow the identification of previously unknown polymorphisms and, ultimately, identification of a gene or pathway involved in the onset and progression of disease.

In addition, the database can be used for taking an identified polymorphism and ascertaining whether it changes in frequency when the data is sorted according to a selected parameter.

- In this way, the databases and methods provided herein permit, among
- 5 other things, identification of components, particularly key components, of a disease process by understanding its genetic underpinnings, and also an understanding of processes, such as individual drug responses. The databases and methods provided herein also can be used in methods involving elucidation of pathological pathways, in developing new diagnostic assays, identifying new
- 10 potential drug targets, and in identifying new drug candidates.

Morbidity and/or early mortality associated polymorphisms

- A database containing information provided by a population of healthy blood donors who were not selected for any particular disease to can be used to identify polymorphisms and the alleles in which they are present, whose
- 15 frequency decreases with age. These may represent morbidity susceptibility markers and genes.

- Polymorphisms of the genome can lead to altered gene function, protein function or genome instability. To identify those polymorphisms which have a clinical relevance/utility is the goal of a world-wide scientific effort. It can be
- 20 expected that the discovery of such polymorphisms will have a fundamental impact on the identification and development of novel drug compounds to cure diseases. However, the strategy to identify valuable polymorphisms is cumbersome and dependent upon the availability of many large patient and control cohorts to show disease association. In particular, genes that cause a
- 25 general risk of the population to suffer from any disease (morbidity susceptibility genes) will escape these case/control studies entirely.

- Here described is a screening strategy to identify morbidity susceptibility genes underlying a variety of different diseases. The definition of a morbidity susceptibility gene is a gene that is expressed in many different cell types or
- 30 tissues (housekeeping gene) and its altered function can facilitate the expression of a clinical phenotype caused by disease-specific susceptibility genes that are involved in a pathway specific for this disorder. In other words, these morbidity

-39-

susceptibility genes predispose people to develop a distinct disease according to their genetic make-up for this disease.

Candidates for morbidity susceptibility genes can be found at the bottom level of pathways involving transcription, translation, heat-shock proteins, 5 protein trafficking, DNA repair, assembly systems for subcellular structures (e.g. mitochondria, peroxysomes and other cellular microbodies), receptor signaling cascades, immunology, etc. Those pathways control the quality of life at the cellular level as well as for the entire organism. Mutations/polymorphisms located in genes encoding proteins for those pathways can reduce the fitness of 10 cells and make the organism more susceptible to express the clinical phenotype caused by the action of a disease-specific susceptibility gene. Therefore, these morbidity susceptibility genes can be potentially involved in a whole variety of different complex diseases if not in all. Disease-specific susceptibility genes are involved in pathways that can be considered as disease-specific pathways like 15 glucose-, lipid, hormone metabolism, etc.

The exemplified method permit, among other things, identification of genes and/or gene products involved in a man's general susceptibility to morbidity and/or mortality; use of these genes and/or gene products in studies to elucidate the genetic underpinnings of human diseases; use of these genes 20 and/or gene products in combinatorial statistical analyses without or together with disease-specific susceptibility genes; use of these genes and/or gene products to predict penetrance of disease susceptibility genes; use of these genes and/or gene products in predisposition and/or acute medical diagnostics and use of these genes and/or gene products to develop drugs to cure diseases 25 and/or to extend the life span of humans.

SCREENING PROCESS

The healthy population stratified by age, gender and ethnicity, etc. is a very efficient and a universal screening tool for morbidity associated genes. Changes of allelic frequencies in the young compared to the old population are 30 expected to indicate putative morbidity susceptibility genes. Individual samples of this healthy population base can be pooled to further increase the throughput. In a proof of principle experiment pools of young and old Caucasian females and

-40-

males were applied to screen more than 400 randomly chosen single nucleotide polymorphisms located in many different genes. Candidate polymorphisms were identified if the allelic difference was greater than 8% between young and old for both or only one of the genders. The initial results were assayed again in at least one independent subsequent experiments. Repeated experiments are necessary to recognize unstable biochemical reactions, which occur with a frequency of about 2-3% and can mimic age-related allelic frequency differences. Average frequency differences and standard deviations are calculated after successful reproducibility of initial results. The final allelic frequency is then compared to a reference population of Caucasian CEPH sample pool. The result should show similar allelic frequencies in the young Caucasian population. Subsequently, the exact allele frequencies of candidates including genotype information were obtained by analyzing all individual samples. This procedure is straight forward with regard to time and cost. It enables the screening of an enormous number of SNPs. So far, several markers with a highly significant association to age were identified and described below.

In general at least 5 individual in a stratified population need to be screened to produce statistically significant results. The frequency of the allele is determined for an age stratified population. Chi square analysis is then performed on the allelic frequencies to determine if the difference between age groups is statistically significant. A p value less than of 0.1 is considered to represent a statistically significant difference. More preferably the p value should be less than 0.05.

Clinical Trials

The identification of markers whose frequency in a population decreases with age also allows for better designed and balanced clinical trials. Currently, if a clinical trial utilizes a marker as a significant endpoint in a study and the marker disappears with age, then the results of the study may be inaccurate. By using methods provided herein, it can be ascertained that if a marker decreases in frequency with age. This information considered and controlled when designing the study. For, example, an age independent marker could be substituted in its place.

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLE 1

This example describes the use of a database containing information 5 provided by a population of healthy blood donors who were not selected for any particular disease to determine the distribution of allelic frequencies of known genetic markers with age and by sex in a Caucasian subpopulation of the database. The results described in this example demonstrate that a disease-related genetic marker or polymorphism can be identified by sorting a healthy 10 database by a parameter or parameters, such as age, sex and ethnicity.

Generating a database

Blood was obtained by venous puncture from human subjects who met blood bank criteria for donating blood. The blood samples were preserved with EDTA at pH 8.0 and labeled. Each donor provided information such as age, sex, 15 ethnicity, medical history and family medical history. Each sample was labeled with a barcode representing identifying information. A database was generated by entering, for each donor, the subject identifier and information corresponding to that subject into the memory of a computer storage medium using commercially available software, e.g., Microsoft Access.

Model genetic markers

The frequencies of polymorphisms known to be associated at some level with disease were determined in a subpopulation of the subjects represented in the database. These known polymorphisms occur in the p21, p53 and Lipoprotein Lipase genes. Specifically, the N291S polymorphism (N291S) of the 25 Lipoprotein Lipase gene, which results in a substitution of a serine for an asparagine at amino acid codon 291, leads to reduced levels of high density lipoprotein cholesterol (HDL-C) that is associated with an increased risk of males for arteriosclerosis and in particular myocardial infarction (see, Reymer *et al.* (1995) *Nature Genetics* 10:28-34).

30 The p53 gene encodes a cell cycle control protein that assesses DNA damage and acts as a transcription factor regulating genes that control cell growth, DNA repair and apoptosis (programmed cell death). Mutations in the

-42-

p53 gene have been found in a wide variety of different cancers, including different types of leukemia, with varying frequency. The loss of normal p53 function results in genomic instability an uncontrolled cell growth. A polymorphism that has been identified in the p53 gene, i.e., the R72P allele,

5 results in the substitution of a proline for an arginine at amino acid codon 72 of the gene.

The p21 gene encodes a cyclin-dependent kinase inhibitor associated with G1 phase arrest of normal cells. Expression of the p21 gene triggers apoptosis. Polymorphisms of the p21 gene have been associated with Wilms' tumor, a pediatric kidney cancer. One polymorphism of the p21 gene, the S31R polymorphism, results in a substitution of an arginine for a serine at amino acid codon 31.

Database analysis

Sorting of subjects according to specific parameters

15 The genetic polymorphisms were profiled within segments of the Caucasian subpopulation of the sample bank. For p53 profiling, the genomic DNA isolated from blood from a total of 1277 Caucasian subjects age 18-59 years and 457 Caucasian subjects age 60-79 years was analyzed. For p21 profiling, the genomic DNA isolated from blood from a total of 910 Caucasian subjects age 18-49 years and 824 Caucasian subjects age 50-79 years was analyzed. For lipoprotein lipase gene profiling, the genomic DNA from a total of 1464 Caucasian females and 1470 Caucasian males under 60 years of age and a total of 478 Caucasian females and 560 Caucasian males over 60 years of age was analyzed.

20 25 Isolation and analysis of genomic DNA

Genomic DNA was isolated from blood samples obtained from the individuals. Ten milliliters of whole blood from each individual was centrifuged at 2000 x g. One milliliter of the buffy coat was added to 9 ml of 155 mM NH₄Cl, 10 mM KHCO₃, and 0.1 mM Na₂EDTA, incubated 10 min at room temperature and centrifuged for 10 min at 2000 x g. The supernatant was removed, and the white cell pellet was washed in 155 mM NH₄Cl, 10 mM KHCO₃ and 0.1 mM Na₂EDTA and resuspended in 4.5 ml of 50 mM Tris, 5 mM

-43-

EDTA and 1% SDS. Proteins were precipitated from the cell lysate by 6 mM ammonium acetate, pH 7.3, and then separated from the nucleic acids by centrifugation at 3000 x g. The nucleic acid was recovered from the supernatant by the addition of an equal volume of 100% isopropanol and
5 centrifugation at 2000 x g. The dried nucleic acid pellet was hydrated in 10 mM Tris, pH 7.6, and 1 mM Na₂EDTA and stored at 4° C.

Assays of the genomic DNA to determine the presence or absence of the known genetic markers were developed using the BiomassPROBE™ detection method (primer oligo base extension) reaction. This method uses a single
10 detection primer followed by an oligonucleotide extension step to give products, which can be readily resolved by mass spectrometry, and, in particular, MALDI-TOF mass spectrometry. The products differ in length depending on the presence or absence of a polymorphism. In this method, a detection primer anneals adjacent to the site of a variable nucleotide or sequence of nucleotides
15 and the primer is extended using a DNA polymerase in the presence of one or more dideoxyNTPs and, optionally, one or more deoxyNTPs. The resulting products are resolved by MALDI-TOF mass spectrometry. The mass of the products as measured by MALDI-TOF mass spectrometry makes possible the determination of the nucleotide(s) present at the variable site.
20 First, each of the Caucasian genomic DNA samples was subjected to nucleic acid amplification using primers corresponding to sites 5' and 3' of the polymorphic sites of the p21 (S31R allele), p53 (R72P allele) and Lipoprotein Lipase (N291S allele) genes. One primer in each primer pair was biotinylated to permit immobilization of the amplification product to a solid support.
25 Specifically, the polymerase chain reaction primers used for amplification of the relevant segments of the p21, p53 and lipoprotein lipase genes are shown below: US4p21c31-2F (SEQ ID NO: 9) and US5p21-2R (SEQ ID NO: 10) for p21 gene amplification; US4-p53-ex4-F (also shown as p53-ex4US4 (SEQ ID NO: 2)) and US5-p53/2-4R (also shown as US5P53/4R (SEQ ID NO: 3)) for p53 gene
30 amplification; and US4-LPL-F2 (SEQ ID NO: 16) and US5-LPL-R2 (SEQ ID NO: 17) for lipoprotein lipase gene amplification.

-44-

Amplification of the respective DNA sequences was conducted according to standard protocols. For example, primers may be used in a concentration of 8 pmol. The reaction mixture (e.g., total volume 50 μ l) may contain 5 Taq-polymerase including 10x buffer and dTNPs. Cycling conditions for polymerase chain reaction amplification may typically be initially 5 min. at 95°C, followed by 1 min. at 94°C, 45 sec at 53°C, and 30 sec at 72°C for 40 cycles with a final extension time of 5 min at 72°C. Amplification products may be purified by using Qiagen's PCR purification kit (No. 28106) according to manufacturer's instructions. The elution of the purified products from the 10 column can be done in 50 μ l TE-buffer (10mM Tris, 1 mM EDTA, pH 7.5).

The purified amplification products were immobilized via a biotin-avidin linkage to streptavidin-coated beads and the double-stranded DNA was denatured. A detection primer was then annealed to the immobilized DNA using conditions such as, for example, the following: 50 μ l annealing buffer (20 mM 15 Tris, 10 mM KCl, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 2 mM MgSO₄, 1% Triton X-100, pH 8) at 50°C for 10 min, followed by washing of the beads three times with 200 μ l washing buffer (40 mM Tris, 1 mM EDTA, 50 mM NaCl, 0.1% Tween 20, pH 8.8) and once in 200 μ l TE buffer.

The PROBE extension reaction was performed, for example, by using 20 some components of the DNA sequencing kit from USB (No. 70770) and dNTPs or ddNTPs from Pharmacia. An exemplary protocol could include a total reaction volume of 45 μ l, containing of 21 μ l water, 6 μ l Sequenase-buffer, 3 μ l 10 mM DTT solution, 4.5 μ l, 0.5 mM of three dNTPs, 4.5 μ l, 2 mM the missing one 25 ddNTP, 5.5 μ l glycerol enzyme dilution buffer, 0.25 μ l Sequenase 2.0, and 0.25 M pyrophosphatase. The reaction can then be pipetted on ice and incubated for 15 min at room temperature and for 5 min at 37°C. The beads may be washed three times with 200 μ l washing buffer and once with 60 μ l of a 70 mM NH₄-Citrate solution.

The DNA was denatured to release the extended primers from the 30 immobilized template. Each of the resulting extension products was separately analyzed by MALDI-TOF mass spectrometry using 3-hydroxypicolinic acid (3-HPA) as matrix and a UV laser.

-45-

Specifically, the primers used in the PROBE reactions are as shown below: P21/31-3 (SEQ ID NO: 12) for PROBE analysis of the p21 polymorphic site; P53/72 (SEQ ID NO: 4) for PROBE analysis of the p53 polymorphic site; and LPL-2 for PROBE analysis of the lipoprotein lipase gene polymorphic site. In 5 the PROBE analysis of the p21 polymorphic site, the extension reaction was performed using dideoxy-C. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 31 encodes a serine) and from the reaction conducted on a polymorphic S31R allele template (wherein codon 31 encodes an arginine) are shown below and designated as P21/31-3 Ser (wt) 10 (SEQ ID NO: 13) and P21/31-3 Arg (SEQ ID NO: 14), respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 4900.2 Da for the wild-type product and 5213.4 Da for the polymorphic product).

In the PROBE analysis of the p53 polymorphic site, the extension reaction 15 was performed using dideoxy-C. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 72 encodes an arginine) and from the reaction conducted on a polymorphic R72P allele template (wherein codon 72 encodes a proline) are shown below and designated as Cod72 G Arg (wt) and Cod72 C Pro, respectively. The masses for each product 20 as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 5734.8 Da for the wild-type product and 5405.6 Da for the polymorphic product).

In the PROBE analysis of the lipoprotein lipase gene polymorphic site, the extension reaction was performed using a mixture of ddA and ddT. The 25 products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 291 encodes an asparagine) and from the reaction conducted on a polymorphic N291S allele template (wherein codon 291 encodes a serine) are shown below and designated as 291Asn and 291Ser, respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also 30 provided (i.e., 6438.2 Da for the wild-type product and 6758.4 Da for the polymorphic product).

P53-1 (R72P)

-46-

PCR Product length: 407 bp (SEQ ID NO: 1)

US4 -p53 -ex4 -F
ctg aggacacctggc cctctgactg

5 ctctttcac ccatctacaq tcccccttgc cgtcccaagc aatqqatqat ttqatqctqt
ccccggacga tattgaacaa tggttcactg aagaccagg tccagatgaa gctcccagaa
P53/72 72R
tgccaaqaggc tgcccccgc gtggccctg caccagcagc tcctacaccg gggccccctg
c 72P

10 caccagcccc ctccctggccc ctgtcatctt ctgtcccttc ccagaaaacc taccaggggca
gctacgttt ccgtctgggc ttcttgatt ctgggacagc caagtctgtg acttgcacgg
tcagttcccc tgaggggctg gcttccatga gacttcaa
US5-p53/2-4R

Primers (SEQ ID NOs: 2-4)

p53-ex4FUS4 ccc agt cac gac qtt gta aaa cgc tga gga cct ggt cct ctg ac
15 US5P53/4R agc gga taa caa ttt cac aca qgt tga agt ctc atg gaa gcc
P53/72 gcc aga ggc tgc tcc cc

Masses

Allele	Product Termination: ddC	SEQ #	Length	Mass
P53/72	gccagaggctgctcccc	5	17	5132.4
20 Cod72 G Arg (wt)	gccagaggctgctcccccgc	6	19	5734.8
Cod72 C Pro	gccagaggctgctcccccc	7	18	5405.6

Biotinylated US5 primer is used in the PCR amplification.

LPL-1 (N291S)

25 Amino acid exchange asparagine to serine at codon 291 of the
lipoprotein lipase gene.

PCR Product length: 251 bp (SEQ ID NO: 15)

US4-LPL-F2 (SEQ ID NO: 16)

30 gcgctccatt catctttca tcgactctct gttgaatgaa gaaaatccaa gtaaggccta
caggtgcagt tccaaggaaag ctttgagaa aggctctgc ttgagttgta gaaagaaccg
LPL-2 291N
ctgcaacaat ctqqqctataq aqatcaataa agtcagagcc aaaagaagca gcaaatgt
g 291S
cctgaagact cgtttcaga tgccc
35 US4-LPL-R2

Primers (SEQ ID NOs: 16-18):

US4-LPL-F2 ccc agt cac gac qtt gta aaa cgg cgc tcc att cat ctc ttc

US5-LPL-R2 agc gga taa caa ttt cac aca qgg ggc atc tga gaa cga gtc

LPL-2 caa tct ggg cta tga gat ca

-47-

Masses

Allele	Product Termination: ddA, ddT	SEQ #	Length	Mass
LPL-2	caatctggctatgagatca	19	20	6141
291 Asn	caatctggctatgagatcaa	20	21	6438.2
5 291 Ser	caatctggctatgagatcagt	21	22	6758.4

Biotinylated US5 primer is used in the PCR amplification.

P21-1 (S31R)

Amino acid exchange serine to arginine at codon 31 of the tumor

- 10 suppressor gene p21. Product length: 207 bp (SEQ ID NO: 8)

US4p21c31-2F

gtcc gtcagaaccc atgcggcagc
p21/31-3 31S

aaggcctgcc gccgcctctt cggcccagtg gacaqcqaqc aqctqaqccg cgactgtgat
15 gcgctaattgg cgggctgcat ccaggaggcc cgtgagcgt ggaacttcga ctttgtcacc
gagacaccac tggaggg
US5p21-2R

Primers (SEQ ID NOs: 9-11)

- 20 US4p21c31-2F ccc agt cac qac gtt qta aaa cg tcc gtc aga acc cat gcg g

US5p21-2R agc qga taa caa ttt cac aca qgc tcc agt ggt gtc tcg gtg ac

P21/31-3 cag cga gca gct gag

Masses

Allele	Product Termination: ddC	SEQ #	Length	Mass
p21/31-3	cagcgagcagctgag	12	15	4627
P21/31-3 Ser (wt)	cagcgagcagctgagc	13	16	4900.2
P21/31-3 Arg	cagcgagcagctgagac	14	17	5213.4

Biotinylated US5 primer is used in the PCR amplification.

- 30 Each of the Caucasian subject DNA samples was individually analyzed by MALDI-TOF mass spectrometry to determine the identity of the nucleotide at the polymorphic sites. The genotypic results of each assay can be entered into the database. The results were then sorted according to age and/or sex to determine the distribution of allelic frequencies by age and/or sex. As depicted in the Figures showing

-48-

histograms of the results, in each case, there was a differential distribution of the allelic frequencies of the genetic markers for the p21, p53 and lipoprotein lipase gene polymorphisms.

Figure 8 shows the results of the p21 genetic marker assays 5 reveals a statistically significant decrease (from 13.3% to 9.2%) in the frequency of the heterozygous genotype (S31R) in Caucasians with age (18-49 years of age compared to 50-79 years of age). The frequencies of the homozygous (S31 and R31) genotypes for the two age groups are also shown, as are the overall frequencies of the S31 and R31 alleles in 10 the two age groups (designated as 'S31 and 'R31, respectively in the Figure).

Figures 7A-C shows the results of the p53 genetic marker assays and reveals a statistically significant decrease (from 6.7% to 3.7%) in the frequency of the homozygous polymorphic genotype (P72) in Caucasians 15 with age (18-59 years of age compared to 60-79 years of age). The frequencies of the homozygous "wild-type" genotype (R72) and the heterozygous genotype (R72P) for the two age groups are also shown, as are the overall frequencies of the R72 and P72 alleles in the two age groups (designated as 'R72 and 'P72, respectively in the Figure). These 20 results are consistent with the observation that allele is not benign, as p53 regulates expression of a second protein, p21, which inhibits cyclin-dependent kinases (CDKs) needed to drive cells through the cell-cycle (a mutation in either gene can disrupt the cell cycle leading to increased cell division).

25 Figure 2C shows the results of the lipoprotein lipase gene genetic marker assays reveals a statistically significant decrease (from 1.97% to 0.54%) in the frequency of the polymorphic allele (S291) in Caucasian males with age (see also Reymer *et al.* (1995) *Nature Genetics* 10:28-34).

-49-

The frequencies of this allele in Caucasian females of different age groups are also shown.

EXAMPLE 2

This example describes the use of MALDI-TOF mass spectrometry

5 to analyze DNA samples of a number of subjects as individual samples and as pooled samples of multiple subjects to assess the presence or absence of a polymorphic allele (the 353Q allele) of the Factor VII gene and determine the frequency of the allele in the group of subjects. The results of this study show that essentially the same allelic frequency can

10 be obtained by analyzing pooled DNA samples as by analyzing each sample separately and thereby demonstrate the quantitative nature of MALDI-TOF mass spectrometry in the analysis of nucleic acids.

Factor VII

Factor VII is a serine protease involved in the extrinsic blood

15 coagulation cascade. This factor is activated by thrombin and works with tissue factor (Factor III) in the processing of Factor X to Factor Xa. There is evidence that supports an association between polymorphisms in the Factor VII gene and increased Factor VII activity which can result in an elevated risk of ischemic cardiovascular disease, including myocardial

20 infarction. The polymorphism investigated in this study is R353Q (i.e., a substitution of a glutamic acid residue for an arginine residue at codon 353 of the Factor VII gene) (see Table 5).

Analysis of DNA samples for the presence or absence of the 353Q allele of the Factor VII gene

25 Genomic DNA was isolated from separate blood samples obtained from a large number of subjects divided into multiple groups of 92 subjects per group. Each sample of genomic DNA was analyzed using the BiomassPROBE™ assay as described in Example 1 to determine the

30 presence or absence of the 353Q polymorphism of the Factor VII gene.

-50-

First, DNA from each sample was amplified in a polymerase chain reaction using primers F7-353FUS4 (SEQ ID NO: 24) and F7-353RUS5 (SEQ ID NO: 26) as shown below and using standard conditions, for example, as described in Example 1. One of the primers was biotinylated

5 to permit immobilization of the amplification product to a solid support. The purified amplification products were immobilized via a biotin-avidin linkage to streptavidin-coated beads and the double-stranded DNA was denatured. A detection primer was then annealed to the immobilized DNA using conditions such as, for example, described in Example 1. The

10 detection primer is shown as F7-353-P (SEQ ID NO: 27) below. The PROBE extension reaction was carried out using conditions, for example, such as those described in Example 1. The reaction was performed using ddG.

The DNA was denatured to release the extended primers from the

15 immobilized template. Each of the resulting extension products was separately analyzed by MALDI-TOF mass spectrometry. A matrix such as 3-hydroxypicolinic acid (3-HPA) and a UV laser could be used in the MALDI-TOF mass spectrometric analysis. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 353

20 encodes an arginine) and from the reaction conducted on a polymorphic 353Q allele template (wherein codon 353 encodes a glutamic acid) are shown below and designated as 353 CGG and 353 CAG, respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 5646.8 Da for the wild-type product

25 and 5960 Da for the polymorphic product).

The MALDI-TOF mass spectrometric analyses of the PROBE reactions of each DNA sample were first conducted separately on each sample (250 nanograms total concentration of DNA per analysis). The allelic frequency of the 353Q polymorphism in the group of 92 subjects

was calculated based on the number of individual subjects in which it was detected.

Next, the samples from 92 subjects were pooled (250 nanograms total concentration of DNA in which the concentration of any individual 5 DNA is 2.7 nanograms) and the pool of DNA was subjected to MALDI-TOF mass spectrometric analysis. The area under the signal corresponding to the mass of the 353Q polymorphism PROBE extension product in the resulting spectrum was integrated in order to quantitate the amount of DNA present. The ratio of this amount to total DNA was used 10 to determine the allelic frequency of the 353Q polymorphism in the group of subjects. This type of individual sample vs. pooled sample analysis was repeated for numerous different groups of 92 different samples.

The frequencies calculated based on individual MALDI-TOF mass spectrometric analysis of the 92 separate samples of each group of 92 15 are compared to those calculated based on MALDI-TOF mass spectrometric analysis of pools of DNA from 92 samples in Figure 9. These comparisons are shown as "pairs" of bar graphs in the Figure, each pair being labeled as a separate "pool" number, e.g., P1, P16, P2, etc. Thus, for example, for P1, the allelic frequency of the polymorphism 20 calculated by separate analysis of each of the 92 samples was 11.41% and the frequency calculated by analysis of a pool of all of the 92 DNA samples was 12.09%.

The similarity in frequencies calculated by analyzing separate DNA samples individually and by pooling the DNA samples demonstrates that it 25 is possible, through the quantitative nature of MALDI-TOF mass spectrometry, to analyze pooled samples and obtain accurate frequency determinations. The ability to analyze pooled DNA samples significantly reduces the time and costs involved in the use of the non-selected, healthy databases as described herein. It has also been shown that it is

-52-

possible to decrease the DNA concentration of the individual samples in a pooled mixture from 2.7 nanograms to 0.27 nanograms without any change in the quality of the spectrum or the ability to quantitate the amount of sample detected.

5 Factor VII R353Q PROBE Assay

PROBE Assay for cod353 CGG>CAG (Arg>Gln), Exon 9 G>A.

PCR fragment: 134 bp (incl. US tags; SEQ ID Nos. 22 and 23)

Frequency of A allele: Europeans about 0.1, Japanese/Chinese about 0.03-0.05 (Thromb. Haemost. 1995, 73:617-22; Diabetologia 1998,

10 41:760-6):

F7-353FUS4 >

1201 GTGCCGGCTA CTCGGATGGC AGCAAGGACT CCTGCAAGGG GGACAGTGGGA
GGCCCATCG

F7-353-P > A < F7-353RUSS5

15 1261 CCACCCACTA CGGGGCACG TGGTACCTGA CGGGCATCGT CAGCTGGGGC
CAGGGCTGCG

Primers (SEQ ID NOS: 24-26)

Tm⁹*

F7-353FUS4 CCC AGT CAC GAC GTT GTA AAA CGA TGG CAG CAA GGA CTC CTG 64°C

F7-353-P CAC ATG CCA CCC ACT ACC

20 F7-353RUSS5 AGC GGA TAA CAA TTT CAC ACA GGT GAC GAT GCC CGT CAG GTA C 64°C

Masses

Allele	Product Termination: ddG	SEQ #	Length	Mass
F7-353-P	atgccacccactacc	27	18	5333.6
353 CGG	cacatgccacccactaccg	28	19	5646.8
25 353 CAG	cacatgccacccactaccag	29	20	5960
US5-bio bio-	agcggtataacaatttcacacagg	30	23	7648.6

Conclusion

The above examples demonstrate an effect of altered frequency of 30 disease causing genetic factors within the general population.

Interpretation of those results allows prediction of the medical relevance of polymorphic genetic alterations. In addition, conclusions can be drawn with regard to their penetrance, diagnostic specificity, positive predictive value, onset of disease, most appropriate onset of preventive strategies,

-53-

and the general applicability of genetic alterations identified in isolated populations to panmixed populations. Therefore, an age- and sex-stratified population-based sample bank that is ethnically homogenous is a suitable tool for rapid identification and validation of genetic factors regarding their
5 potential medical utility.

EXAMPLE 3

MORBIDITY AND MORTALITY MARKERS

Sample Band and Initial Screening

Healthy samples were obtained through the blood bank of San
10 Bernardino, CA. Donors signed prior to the blood collection a consent form and agreed that their blood will be used in genetic studies with regard to human aging. All samples were anonymized. Tracking back of samples is not possible.

Isolation of DNA from blood samples of a healthy donor population

15 Blood is obtained from a donor by venous puncture and preserved with 1mM EDTA pH 8.0. Ten milliliters of whole blood from each donor was centrifuged at 2000x g. One milliliter of the buffy coat was added to 9 milliliters of 155mM NH₄Cl, 10mM KHCO₃, and 0.1mM Na₂EDTA, incubated 10 minutes at room temperature and centrifuged for 10
20 minutes at 2000x g. The supernatant was removed, and the white cell pellet was washed in 155mM NH₄Cl, 10mM KHCO₃, and 0.1mM Na₂EDTA and resuspended in 4.5 milliliters of 50mM Tris, 5mM EDTA, and 1% SDS. Proteins were precipitated from the cell lysate by 6M Ammonium Acetate, pH 7.3, and separated from the nucleic acid by
25 centrifugation 3000x g. The nucleic acid was recovered from the supernatant by the addition of an equal volume of 100% isopropanol and centrifugation at 2000x g. The dried nucleic acid pellet was hydrated in 10mM Tris pH 7.6 and 1mM Na₂EDTA and stored at 4C.

-54-

In this study, samples were pooled as shown in Table 1. Both parents of the blood donors were of Caucasian origin.

Table 1

	Pool ID	Sex	Age-range	# individuals
5	SP1	Female	18-39 years	276
	SP2	Males	18-39 years	276
	SP3	Females	60-69 years	184
	SP4	Males	60-79 years	368

- 10 More than 400 SNPs were tested using all four pools. After one test run
 34 assays were selected to be re-assayed at least once. Finally, 10
 assays showed repeatedly differences in allele frequencies of several
 percent and, therefore, fulfilled the criteria to be tested using the
 individual samples. Average allele frequency and standard deviation is
 15 tabulated in Table 2.

Table 2

	Assay ID	SP1	SP1-STD	SP2	SP2-STD	SP3	SP3-STD	SP4	SP4-STD
20	47861	0.457	0.028	0.433	0.042	0.384	0.034	0.380	0.015
	47751	0.276	0.007	0.403	0.006	0.428	0.052	0.400	0.097
	48319	0.676	0.013	0.627	0.018	0.755	0.009	0.686	0.034
	48070	0.581	0.034	0.617	0.045	0.561	n.a.	0.539	0.032
	49807	0.504	0.034	0.422	0.020	0.477	0.030	0.556	0.005
	49534	0.537	0.017	0.503	n.a.	0.623	0.023	0.535	0.009
25	49733	0.560	0.006	0.527	0.059	0.546	0.032	0.436	0.016
	49947	0.754	0.008	0.763	0.047	0.736	0.052	0.689	0.025
	50128	0.401	0.022	0.363	0.001	0.294	0.059	0.345	0.013

-55-

63306	0.697	0.012	0.674	0.013	0.712	0.017	0.719	0.005
-------	-------	-------	-------	-------	-------	-------	-------	-------

So far, 7 out of the 10 potential morbidity markers were fully analyzed. Additional information about genes in which these SNPs are located was gathered through publicly databases like Genbank.

AKAPS

Candidate morbidity and mortality markers include housekeeping genes, such as genes involved in signal transduction. Among such genes are the A-kinase anchoring proteins (AKAPs) genes, which participate in signal transduction pathways involving protein phosphorylation. Protein phosphorylation is an important mechanism for enzyme regulation and the transduction of extracellular signals across the cell membrane in eukaryotic cells. A wide variety of cellular substrates, including enzymes, membrane receptors, ion channels and transcription factors, can be phosphorylated in response to extracellular signals that interact with cells. A key enzyme in the phosphorylation of cellular proteins in response to hormones and neurotransmitters is cyclic AMP (cAMP)-dependent protein kinase (PKA). Upon activation by cAMP, PKA thus mediates a variety of cellular responses to such extracellular signals. An array of PKA isozymes are expressed in mammalian cells. The PKAs usually exist as inactive tetramers containing a regulatory (R) subunit dimer and two catalytic (C) subunits. Genes encoding three C subunits (Ca , $\text{C}\beta$ and $\text{C}\gamma$) and four R subunits ($\text{RI}\alpha$, $\text{RI}\beta$, $\text{RII}\alpha$ and $\text{RII}\beta$) have been identified [see Takio *et al.* (1982) *Proc. Natl. Acad. Sci. U.S. A.* 79:2544-2548; Lee *et al.* (1983) *Proc. Natl. Acad. Sci. U.S. A.* 80:3608-3612; Jahnson *et al.* (1996) *J. Biol. Chem.* 261:12352-12361; Clegg *et al.* (1988) *Proc. Natl. Acad. Sci. U.S. A.* 85:3703-3707; and Scott (1991) *Pharmacol. Ther.* 50:123-145]. The type I (RI) α and type II (RII) α subunits are distributed ubiquitously, whereas RI β and RII β are present mainly in brain [see, e.g., Miki and Eddy

(1999) *J. Biol. Chem.* 274:29057-29062]. The type I PKA holoenzyme (RI α and RI β) is predominantly cytoplasmic, whereas the majority of type II PKA (RII α and RII β) associates with cellular structures and organelles [Scott (1991) *Pharmacol. Ther.* 50:123-145]. Many hormones and other signals act through receptors to generate cAMP which binds to the R subunits of PKA and releases and activates the C subunits to phosphorylate proteins. Because protein kinases and their substrates are widely distributed throughout cells, there are mechanisms in place in cells to localize protein kinase-mediated responses to different signals. One such mechanism involves subcellular targeting of PKAs through association with anchoring proteins, referred to as A-kinase anchoring proteins (AKAPs), that place PKAs in close proximity to specific organelles or cytoskeletal components and particular substrates thereby providing for more specific PKA interactions and localized responses [see, e.g., Scott *et al.* (1990) *J. Biol. Chem.* 265:21561-21566; Bregman *et al.* (1991) *J. Biol. Chem.* 266:7207-7213; and Miki and Eddy (1999) *J. Biol. Chem.* 274:29057-29062]. Anchoring not only places the kinase close to preferred substrates, but also positions the PKA holoenzyme at sites where it can optimally respond to fluctuations in the second messenger cAMP [Mochly-Rosen (1995) *Science* 268:247-251; Faux and Scott (1996) *Trends Biochem. Sci.* 21:312-315; Hubbard and Cohen (1993) *Trends Biochem. Sci.* 18:172-177].

Up to 75% of type II PKA is localized to various intracellular sites through association of the regulatory subunit (RII) with AKAPs [see, e.g., Hausken *et al.* (1996) *J. Biol. Chem.* 271:29016-29022]. RII subunits of PKA bind to AKAPs with nanomolar affinity [Carr *et al.* (1992) *J. Biol. Chem.* 267:13376-13382], and many AKAP-RII complexes have been isolated from cell extracts. RI subunits of PKA bind to AKAPs with only micromolar affinity [Burton *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A.*

-57-

94:11067-11072]. Evidence of binding of a PKA RI subunit to an AKAP has been reported [Miki and Eddy (1998) *J. Biol. Chem.* 273:34384-34390] in which RII α -specific and RII α /RII β dual specificity PKA anchoring domains were identified on FSC1/AKAP82. Additional dual specific
5 AKAPs, referred to as D-AKAP1 and D-AKAP2, which interact with the type I and type II regulatory subunits of PKA have also been reported [Huang *et al.* (1997) *J. Biol. Chem.* 272:8057-8064; Huang *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:11184-11189].

More than 20 AKAPs have been reported in different tissues and
10 species. Complementary DNAs (cDNAs) encoding AKAPs have been isolated from diverse species, ranging from *Caenorhabditis elegans* and *Drosophila* to human [see, e.g., Colledge and Scott (1999) *Trends Cell Biol.* 9:216-221]. Regions within AKAPs that mediate association with RII subunits of PKA have been identified. These regions of approximately
15 10-18 amino acid residues vary substantially in primary sequence, but secondary structure predictions indicate that they are likely to form an amphipathic helix with hydrophobic residues aligned along one face of the helix and charged residues along the other [Carr *et al.* (1991) *J. Biol. Chem.* 266:14188-14192; Carr *et al.* (1992) *J. Biol. Chem.* 267:13376-13382]. Hydrophobic amino acids with a long aliphatic side chain, e.g., valine, leucine or isoleucine, may participate in binding to RII subunits
20 [Glantz *et al.* (1993) *J. Biol. Chem.* 268:12796-12804].

Many AKAPs also have the ability to bind to multiple proteins, including other signaling enzymes. For example, AKAP79 binds to PKA,
25 protein kinase C (PKC) and the protein phosphatase calcineurin (PP2B) [Coghlan *et al.* (1995) *Science* 267:108-112 and Klauck *et al.* (1996) *Science* 271:1589-1592]. Therefore, the targeting of AKAP79 to neuronal postsynaptic membranes brings together enzymes with opposite catalytic activities in a single complex.

-58-

AKAPs thus serve as potential regulatory mechanisms that increase the selectivity and intensity of a cAMP-mediated response. There is a need, therefore, to identify and elucidate the structural and functional properties of AKAPs in order to gain a complete understanding of the
5 important role these proteins play in the basic functioning of cells.

AKAP10

The sequence of a human AKAP10 cDNA (also referred to as D-AKAP2) is available in the GenBank database, at accession numbers AF037439 (SEQ ID NO: 31) and NM 007202. The AKAP10 gene is
10 located on chromosome 17.

The sequence of a mouse D-AKAP2 cDNA is also available in the GenBank database (see accession number AF021833). The mouse D-AKAP2 protein contains an RGS domain near the amino terminus that is characteristic of proteins that interact with G α subunits and possess
15 GTPase activating protein-like activity [Huang *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:11184-11189]. The human AKAP10 protein also has sequences homologous to RGS domains. The carboxy-terminal 40 residues of the mouse D-AKAP2 protein are responsible for the interaction with the regulatory subunits of PKA. This sequence is fairly well
20 conserved between the mouse D-AKAP2 and human AKAP10 proteins.

Polymorphisms of the human AKAP10 gene and polymorphic AKAP10 proteins

Polymorphisms of AKAP genes that alter gene expression, regulation, protein structure and/or protein function are more likely to
25 have a significant effect on the regulation of enzyme (particularly PKA) activity, cellular transduction of signals and responses thereto and on the basic functioning of cells than polymorphisms that do not alter gene and/or protein function. Included in the polymorphic AKAPs provided herein are human AKAP10 proteins containing differing amino acid
30 residues at position number 646.

-59-

Amino acid 646 of the human AKAP10 protein is located in the carboxy-terminal region of the protein within a segment that participates in the binding of R-subunits of PKAs. This segment includes the carboxy-terminal 40 amino acids.

- 5 The amino acid residue reported for position 646 of the human AKAP10 protein is an isoleucine. Polymorphic human AKAP10 proteins provided herein have the amino acid sequence but contain residues other than isoleucine at amino acid position 646 of the protein. In particular embodiments of the polymorphic human AKAP10 proteins provided
10 herein, the amino acid at position 646 is a valine, leucine or phenylalanine residue.

An A to G transition at nucleotide 2073 of the human AKAP10 coding sequence

- As described herein, an allele of the human AKAP10 gene that
15 contains a specific polymorphism at position 2073 of the coding sequence and thereby encodes a valine at position 646 has been detected in varying frequencies in DNA samples from younger and older segments of the human population. In this allele, the A at position 2073 of the AKAP10 gene coding sequence is changed from an A to a G, giving rise
20 to an altered sequence in which the codon for amino acid 646 changes from ATT, coding for isoleucine, to GTT, coding for valine.

Morbidity marker 1: human protein kinase A anchoring protein (AKAP10-1)

- PCR Amplification and BiomassPROBE assay detection of AKAP10-1 in a
25 healthy donor population

PCR Amplification of donor population for AKAP 10

- PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10 target sequence was carried out in single 50 μ l PCR reaction with 100ng-1ug of pooled human genomic
30 DNAs in a 50 μ l PCR reaction. Individual DNA concentrations within the

-60-

- pooled samples were present in equal concentration with the final concentration ranging from 1-25ng. Each reaction containing IX PCR buffer (Qiagen, Valencia, CA), 200UμM dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl₂, and 25pmol of the
- 5 forward primer containing the universal primer sequence and the target specific sequence 5'-TCTCAATCATGTGCATTGAGG-3'(SEQ ID NO: 45), 2pmol of the reverse primer
5'-AGCGGATAACAATTTCACACAGGGATCACACAGCCATCAGCAG-3'
(SEQ ID NO: 46), and 10pmol of a biotinylated universal primer
- 10 complementary to the 5' end of the PCR amplicon
5'-AGCGGATAACAATTTCACACAGG-3'(SEQ ID NO: 47). After an initial round of amplification with the target with the specific forward and reverse primer, the 5' biotinylated universal primer then hybridized and acted as a reverse primer thereby introducing a 3' biotin capture moiety
- 15 into the molecule. The amplification protocol results in a 5'-biotinylated double stranded DNA amplicon and dramatically reduces the cost of high throughput genotyping by eliminating the need to 5' biotin label each forward primer used in a genotyping. Thermal cycling was performed in 0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler
- 20 (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles: 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec; 72° C 3min.

Immobilization of DNA

- The 50μl PCR reaction was added to 25ul of streptavidin coated magnetic
- 25 bead (Dynal) prewashed three times and resuspended in 1M NH₄Cl, 0.06M NH₄OH. The PCR amplicons were allowed to bind to the beads for 15 minutes at room temperature. The beads were then collected with a magnet and the supernatant containing unbound DNA was removed. The unbound strand was release from the double stranded amplicons by

-61-

incubation in 100mM NaOH and washing of the beads three times with 10mM Tris pH 8.0.

BiomassPROBE assay analysis of donor population for AKAP10-1 (clone 48319)

- 5 Genotyping using the BiomassPROBE assay methods was carried out by resuspending the DNA coated magnetic beads in 26mM Tris-HCl pH 9.5, 6.5 mM MgCl₂ and 50mM each of dTTP and 50mM each of ddCTP, ddATP, ddGTP, 2.5U of a thermostable DNA polymerase (Ambersham) and 20pmol of a template specific oligonucleotide PROBE
- 10 primer 5'-CTGGCGCCCACGTGGTCAA-3' (SEQ ID NO: 48) (Operon). Primer extension occurs with three cycles of oligonucleotide primer hybridization and extension. The extension products were analyzed after denaturation from the template with 50mM NH₄Cl and transfer of 150nL each sample to a silicon chip preloaded with 150nL of H3PA matrix
- 15 material. The sample material was allowed to crystallize and was analyzed by MALDI-TOF (Bruker, PerSeptive). The SNP that is present in AKAP10-1 is a T to C transversion at nucleotide number 156277 of the sequence of a genomic clone of the AKAP10 gene (GenBank Accession No. AC005730) (SEQ ID NO: 36). SEQ ID NO: 35: represents the
- 20 nucleotide sequence of human chromosome 17, which contains the genomic nucleotide sequence of the human AKAP10 gene, and SEQ ID NO: represents the nucleotide sequence of human chromosome 17, which contains the genomic nucleotide sequence of the human AKAP10-1 allele. The mass of the primer used in the BioMass probe reaction was
- 25 5500.6 daltons. In the presence of the SNP, the primer is extended by the addition of ddC, which has a mass of 5773.8. The wildtype gene results in the addition of dT and ddG to the primer to produce an extension product having a mass of 6101 daltons.

-62-

The frequency of the SNP was measured in a population of age selected healthy individuals. Five hundred fifty-two (552) individuals between the ages of 18-39 years (276 females, 276 males) and 552 individuals between the ages of 60-79 (184 females between the ages of 5 60-69, 368 males between the age of 60-79) were tested for the presence of the polymorphism localized in the non-translated 3'region of AKAP 10. Differences in the frequency of this polymorphism with increasing age groups were observed among healthy individuals.

Statistical analysis showed that the significance level for differences in 10 the allelic frequency for alleles between the "younger" and the "older" populations was $p = 0.0009$ and for genotypes was $p = 0.003$. Differences between age groups are significant. For the total population allele significance is $p = 0.0009$, and genotype significance is $p = 0.003$.

This marker led to the best significant result with regard to allele 15 and genotype frequencies in the age-stratified population. Figure 19 shows the allele and genotype frequency in both genders as well as in the entire population. For latter the significance for alleles was $p = 0.0009$ and for genotypes was $p = 0.003$. The young and old populations were in Hardy-Weinberg equilibrium. A preferential change of one particular 20 genotype was not seen.

The polymorphism is localized in the non-translated 3'-region of the gene encoding the human protein kinase A anchoring protein (AKAP10). The gene is located on chromosome 17. Its structure includes 15 exons and 14 intervening sequences (introns). The encoded protein is 25 responsible for the sub-cellular localization of the cAMP-dependent protein kinase and, therefore, plays a key role in the G-protein mediated receptor-signaling pathway (Huang et al. PNAS (1007) 94:11184-11189). Since its localization is outside the coding region, this polymorphism is most likely in linkage disequilibrium (LD) with other non-synonymous

-63-

polymorphisms that could cause amino acid substitutions and subsequently alter the function of the protein. Sequence comparison of different Genbank database entries concerning this gene revealed further six potential polymorphisms of which two are supposed to change the 5 respective amino acid (see Table 3).

Table 3

Exon	Codon	Nucleotides	Amino acid
10	3 100	GCT > GCC	Ala > Ala
	4 177	AGT > GTG	Met > Val
	8 424	GGG > GGC	Gly > Gly
	10 524	CCG > CTG	Pro > Leu
	12 591	GTG > GTC	Val > Val
	12 599	CGC > CGA	Arg > Arg

15 **Morbidity marker 2: human protein kinase A anchoring protein
(AKAP10-5)**

Discovery of AKAP10-5 Allele (SEQ ID NO: 33)

Genomic DNA was isolated from blood (as described above) of seventeen (17) individuals with a genotype CC at the AKAP10-1 gene 20 locus and a single heterozygous individual (CT) (as described). A target sequence in the AKAP10-1 gene which encodes the C-terminal PKA binding domain was amplified using the polymerase chain reaction. PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10-1 target sequence was carried out in 25 individual 50µl PCR reaction with 25ng of human genomic DNA templates. Each reaction containing 1 X PCR buffer (Qiagen, Valencia, CA), 200µM dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl₂, 25pmol of the forward primer (Ex13F) containing the universal primer sequence and the target specific sequence 5'-TCC CAA 30 AGT GCT GGA ATT AC-3' (SEQ ID NO: 53), and 2pmol of the reverse

54). Thermal cycling was performed in 0.2 mL tubes or 96 well plate

using an MJ Research Thermal Cycler (MJ Research, Waltham, MA)

(calculated temperature) with the following cycling parameters: 94° C for

5 min; 45 cycles; 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec;

72° C 3 min. After amplification the amplicons were purified using a

chromatography (MJO Bio Laboratories (Sollana Beach, CA)).

The sequence of the 18 amplicons, representing the target region,

was determined using a standard Sanger cycle sequencing method with

25 nmol of the PCR amplicon, 3.2 μM DNA sequencing primer 5'-CCG ACA

GCA GTT AAT CCT TC-3' (SEQ ID NO: 55), and chain terminating

drhodamine labeled 2', 3', diideoxynucleotides (PE Biosystems, Foster

City, CA) using the following cycling parameters: 96° C for 15 seconds;

25 cycles: 55° C for 15 seconds, 60° C for 4 minutes. The sequencing

products precipitated by 0.3 M NaOAc and ethanol. The precipitate was

centrifuged and dried. The pellets were resuspended in deionized

formamide and separated on a 5% polyacrylimide gel. The sequence was

determined using the "Sequencher" software (Gene Codes, Ann Arbor,

MI).

20 The sequence of all 17 of the amplicons, which are homozygous

a cDNA clone of the wildtype AKAP10 (GenBank Accession No.

AC005730 for AKAP10 genomic clone (SEQ ID NO: 35)) with A replaced

by G. This SNP can also be designated as located at nucleotide 2073 of

AF037439 (SEQ ID NO: 31). The amino acid sequence of the human

AKAP10 protein is provided as SEQ ID NO: 32. This single nucleotide

polyorphism was designated as AKAP10-5 (SEQ ID NO: 33) and

resulted in a substitution of a valine for an isoleucine residue at amino

residues 331 and 332.

acid position 646 of the amino acid sequence of human AKAP10 (SEQ ID NO: 32). PCR Amplification and BiomassPROBE assay detection of AKAP10-5 in a healthy donor population The healthy population stratified by age is a very efficient and a universal screening tool for morbidity associated genes by allowing for the detection of changes of allelic frequencies in the young compared to the old population. Individual samples of this healthy population base can be pooled to further increase the throughput.

Healthy samples were obtained through the blood bank of San Bernardino, CA. Both parents of the blood donors were of Caucasian origin. Practically a healthy subject, when human, is defined as human viral, bacterial, mycoplasma, and parasitic infections; not anemic; and then further selected based upon a questionnaire regarding history (see Figure 3). Thus, a healthy population represents an unbiased population of sufficient health to donate blood according to blood bank criteria, and not further selected for any disease state. Typically such individuals are in a single 50 μ l PCR reaction with 100ng-1 μ g of pooled human genomic DNA in a 50 μ l PCR reaction. Individual DNA concentrations within the pool of samples were present in equal concentration with the final PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10 target sequence was carried out in a single 50 μ l PCR reaction with 100ng-1 μ g of pooled human genomic DNA in a 50 μ l PCR reaction with 1-25ng. Each reaction contained 1X PCR buffer (Qiagen, Valencia, CA), 4 mM MgCl₂, and 25pmol of the forward primer (Qiagen, Valencia, CA), 200 μ M DNTPs, 1 U Hostar Taq polymerase (Qiagen, Valencia, CA), 200 μ M DNTPs, 1 U Hostar Taq polymerase concentrate ranging from 1-25ng. Each reaction contained 1X PCR containing the universal primer sequence 5'-AGCGGATACACAATTCAACAGGGAGCTTGAAAGAT sequence 5'-AGCGGATACACAATTCAACAGGGAGCTTGAAAGAT

Patent No. 6,043,031) of donor population for AKAP 10-5 (SEQ ID No.: BiomassPROBE_{TM} assay of primer extension analysis (see, U.S.

Detection of AKAP10-5 using BiomassPROBE_{TM} Assay

of the beads three times with 10MM Tris pH 8.0.
25 double stranded amplicons by incubation in 100MM NaOH and washing was removed. The hybridized but unbound strand was released from the collected with a magnet and the supernatant containing unbound DNA beads for 15 minutes at room temperature. The beads were then strand of the double stranded PCR amplicons were allowed to bind to the times and resuspended in 1M NH₄Cl, 0.06M NH₄OH. The 5' end of one 20 magnetic beads (Dynal, Oslo, Norway), which were prewashed three The 50 μl PCR reaction was added to 25μl of streptavidin coated immobilization of DNA

sec, 56° C for 30 sec; 72° C for 60 sec; 72° C 3min.

15 following cycling parameters: 94° C for 5 min; 45 cycles: 94° C for 20 sec, 56° C for 30 sec; 72° C for 60 sec; 72° C 3min.

using an MJ Research Thermal Cycler (calculated temperature) with the thermal cycling was performed in 0.2ml tubes or 96 well plate genotyping.

eliminating the need to 5' biotin label every forward primer used in a 10 dramatically reduced the cost of high throughput genotyping by resulted in a 5'-biotinylated double stranded DNA amplicon and biotin capture moiety into the molecule. The amplification protocol be hybridized and acted as a forward primer thereby introducing a 5' forward and reverse primer, the 5' biotinylated universal primer can then 5 After an initial round of amplification with the target with the specific amplicon BIO:5'-AGCGGATACCAATTCAACAGG-3', (SEQ ID NO: 43). biotinylated universal primer complementary to the 5' end of the PCR 5-GTCACATAATGCAAACAGTG-3', (SEQ ID NO: 54), and 10pmol of a TGC-3', (SEQ ID NO: 41), 2pmol of the reverse primer

33) was performed. Genotyping using these methods was carried out by resuspending the DNA coated magnetic beads in 26mM Tris-HCl pH 9.5, 6.5 mM MgCl₂, 50mM dTTP, 50mM each of ddCTP, ddATP, ddGTP, 2.5U of a thermostable DNA polymerase (AmpliSham), and 20pmol of a template specific oligonucleotide PROBE primer 5'-ACTGAGCCTGCATAA-3' (SEQ ID NO: 44) (Operon). Primer extension occurs with three cycles of oligonucleotide primer with hybridization and extension. The extension products were analyzed after denaturation from the template with 50 mM NH₄Cl and transfer of 150 nL of each sample to a silicon chip preloaded with 150 nL of H3PA matrix material. The sample material was allowed to crystallize and analyzed by MALDI-TOF (Bruker, PerSeptive). The primer has a mass of 5483.6 daltons. The SNP results in the additional of a ddC to the primer, giving a mass of 5756.8 daltons for the extended product. The wild type results in the addition a T and ddG to the primer giving a mass of 6101 daltons. The frequency of the SNP was measured in a population of age under 40 years of age (360 females, 353 males) and 703 individuals over 60 years of age (322 females, 381 males) were tested for the presence of the SNP, AKAP10-5 (SEQ ID NO: 33). Results are presented below in Table 1.

AKAP10-5 (2073V) frequency comparison in 2 age groups					
Female	Alleles	*G	<40	>60	delta G allele
	Genotypes	G	13.9	11.8	2.1
		*A	61.4	65.4	
			38.6	34.6	4.0
			A	36.7	42.5

TABLE 1

Male		All alleles	*G	41.4	37.0	4.4
		Genotypes	G	18.4	10.8	7.7
		GA	45.9	52.5		
		A	35.7	36.7		
Total		All alleles	*G	40.0	35.9	4.1
		Genotypes	G	16.1	11.2	4.9
		GA	47.7	49.4		
		*A	60.0	64.1		
Total		All alleles	*G	36.2	39.4	

Methionine sulfoxide reductase A (#63306) PCR Amplification and BiomassPROBE assay detection of the human methionine sulfoxidase A (h-msr-A) in a healthy donor population PCR Amplification of donor population for h-msr-A PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10 target sequence was carried out in single 50 μ l PCR reaction with 100ng-1ug of pooled human genomic DNA template in a 50 μ l PCR reaction. Individual DNA concentrations within the pooled samples were present in an equal concentration with highly significant.

Figure 20 graphically shows these results of allele and genotype distribution in the age and sex stratified Caucasian population. Morbidity marker 3: human methionine sulfoxide reductase A (msra) The age-related allele and genotype frequency of this marker in both genders and the entire population is shown in Figure 21. The decrease of the homozygous CC genotype in the older male population is highly significant.

Male		All alleles	*G	41.4	37.0	4.4		
		All alleles	*A	58.6	63.0			
		Genotypes	G	18.4	10.8	7.7		
		GA	45.9	52.5	36.7			
		A	35.7					
		Alleles	*G	40.0	35.9	4.1		
Total		Genotypes	G	16.1	11.2	4.9		
		GA	47.7	49.4				
		A	36.2	39.4				

the final concentration ranging from 1-25ng. Each reaction containing 1X PCR buffer (Qiagen, Valencia, CA), 200 μ M DNTPs, 1U Hostar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl₂, 25pmol of the forward primer containing the sequence 5'-TTCTCTGCCACAGAGAGC-3'. (SEQ ID NO: 50), and 10pmol of a biotinylated universal primer (SEQ ID NO: 50). After an initial round of amplification of the target with the specific forward and reverse complementary to the 5' end of the PCR amplicon sequence 5'-AGCGGATACACAATTCAACAGG-3'. (SEQ ID NO: 51). A reverse transcriptase, the 5', biotinylated universal primer was then hybridized and acted as a reverse primer thereby introducing a 3', biotin capture moiety into the molecule. The amplification protocol results in a 5'-biotinylated double stranded DNA amplicon and dramatically reduces the cost of forward primer used in a genotyping. Thermal cycling was performed in 0.2ml tubes or 96 well plate using an MJ Research Thermal Cycler 15 high throughput genotyping by eliminating the need to 5', biotin label each 0.2ml tube. The calculated temperature with the following cycling parameters: 94° C for 45 cycles; 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec; 20 5 min; 45 cycles; 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec; (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles; 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec; 25 72° C 3min.

Immobilization of DNA

The 50 μ l PCR reaction was added to 25 μ l of streptavidin coated magnetic bead (Dynal) prewashed three times and resuspended in 1M NH₄Cl, 0.06M NH₄OH. The PCR amplicons were allowed to bind to the beads for 15 minutes at room temperature. The beads were then collected with a magnet and the supernatant containing unbound DNA was removed. The unbound strand was release from the double stranded

Genotyping using the BiomassPROBE assay methods was carried out by resuspending the he DNA coated magnetic beads in 26mM Tris-HCl pH 9.5, 6.5 M MgCl₂, 50mM of DTTs and 50mM each of ddCTP, ddATP, ddGTP, 2.5U of a thermostable DNA polymerase (Ambersham), and 20pmol of a template specific oligonucleotide PROBE (Ambersham), primer 5'-CTGAAAGGGAGAAG-3' (Operon) (SEQ ID NO: 52). Primer extension occurs with three cycles of oligonucleotide primer with hybridization and extension. The extension products were analyzed after denaturation from the template with 50M NH₄Cl and transfer of 150nl each sample to a silicon chip preloaded with 150nl of H3PA matrix material. The sample material was allowed to crystallize and analyzed by MALDI-TOF (Bruker, PerSeptive). The SNP is represented as a T to C transversion in the sequence of two ESTs. The wild type is represented by having a T at position 128 of GenBank Accession No. AW 195104, which represents the nucleotide sequence of an EST which is a portion of the wild type human msrA gene (SEQ ID NO: 39). The SNP is presented as a C at position 129 of GenBank Accession No. AW 874187, which represents the nucleotide sequence of an EST which is a portion of an allele of the human msrA gene (SEQ ID NO: 40).

In a genomic sequence the SNP is represented as an A to G transversion. The primer utilized in the BioMass probe reaction had a mass of 5654.8 daltons. In the presence of the SNP the primer is extended by the incorporation of a dC and has a mass of 5928. In the presence of the wildtype the primer is extended by adding a dT and a mass of 6232.1 daltons.

The frequency of the SNP was measured in a population of age 60-69, 368 males between the age of 60-79 (1184 females, 276 males and 552 between the ages of 18-39 years (276 females, 276 males and 552 selected healthy individuals. Five hundred fifty-two (552) individuals between the ages of 60-79 were tested for the presence of the polymorphism localized in the non-translated 3'-region of h-msr-A. Genotype difference between male age group among healthy individuals is significant. For the male population allele significance is $p=0.0009$ and genotype significance is $p=0.003$. The age-related allele and genotype frequency of this marker in both genders and the entire population is shown in Figure 21. The decrease of the homozygous CC genotype in the older male population is highly significant.

The polymorphism is localized in the non-translated 3'-region of the gene encoding the human methionine sulfoxide reductase (h-msrA). The exact localization is 451 base pairs downstream the stop codon (TAA). It is very likely that this SNP is in linkage disequilibrium (LD) with another polymorphism more upstream in the coding or promoter region; thus, it is not directly cause morbidity. The enzyme methionine sulfoxide reductase has been proposed to exhibit multiple biological functions. It may serve to repair oxidative protein damage but also play an important role in the regulation of proteins by activation or inactivation of their biological functions (Moscovitz et al. (1990) PNAS 95:14071-14075). It has also been shown that its activity is significantly reduced in brain tissues of Alzheimer patients (Gabbita et al. (1999) J. Neurochem 73:1660-1666).

It is scientifically conceivable that proteins involved in the metabolism of reactive oxygen species are associated to disease.

CONCLUSION

-72-

All of the products of the enzyme assays listed below were analyzed by MALDI-TOF mass spectrometry. A diluted matrix solution (0.15 μ L) containing of 10:1 3-hydroxyprocolinic acid:ammonium citrate in 1:1 water:acetonitrile diluted 2.5-fold with water was pipetted onto a SpectroChip (Sequenom, Inc.) and was allowed to crystallize. Then, 0.15 μ L of sample was added. A linear Positive Voyager DE mass spectrometer or Bruker BiFlex MALDI-TOF mass spectrometer, operating in positive ion mode, was used for the measurements. The sample plates were kept at 18.2 KV for 400 nm after each UV laser shot (approximate 250 laser shots total), and then the target voltage was raised to 20 KV. The original spectra were digitized at 500 MHz.

MALDI-TOF Mass Spectrometry Analysis

EXAMPLE 4

The use of the healthy population provides for the identification of morbidity markers. The identification of proteins involved in the G-protein coupled signaling transduction pathway or in the detoxification of oxidative stress can be considered as convincing results. Further confirmation and validation of other potential polymorphisms already identified in *silico* in the gene encoding the human protein kinase A anchoring protein could even provide stronger association to morbidity and demonstrate that this gene product is a suitable pharmaceutical diagnostic target.

Where indicated in the examples below, the products of the enzymatic digestions were purified with ZipTips (Millipore, Bedford, MA). The ZipTips were pre-wetted with 10 μ L 50% acetonitrile and equilibrated 4 times with 10 μ L 0.1 M TEAC. The oligonucleotide fragments were

25 Sample Conditioning

EXAMPLE 5

20 All of the products of the enzyme assays listed below were analyzed by MALDI-TOF mass spectrometry. A diluted matrix solution (0.15 μ L) containing of 10:1 3-hydroxyprocolinic acid:ammonium citrate in 1:1 water:acetonitrile diluted 2.5-fold with water was pipetted onto a SpectroChip (Sequenom, Inc.) and was allowed to crystallize. Then, 0.15 μ L of sample was added. A linear Positive Voyager DE mass spectrometer or Bruker BiFlex MALDI-TOF mass spectrometer, operating in positive ion mode, was used for the measurements. The sample plates were kept at 18.2 KV for 400 nm after each UV laser shot (approximate 250 laser shots total), and then the target voltage was raised to 20 KV. The original spectra were digitized at 500 MHz.

bound to the C18 in the ZipTip material by continuous aspiration and dispersion of each sample into the ZipTip. Each digested oligonucleotide was conditioned by washing with 10 μ L 0.1 M TEAAc, followed by 4 washing steps with 10 μ L H₂O. DNA fragments were eluted from the

5 Ziptip with 7 μ L 50% acetonitrile.

Any method for condition the samples may be employed. Methods for conditioning, which generally is used to increase peak resolution, are well known (see, e.g., International PCT application No. WO 98/20019).

EXAMPLE 6

10 **DNA Glycosylase-Mediated Sequence Analysis**

DNA Glycosylases modifies DNA at each position that a specific nucleobase resides in the DNA, thereby producing abasic sites. In a subsequent reaction with another enzyme, a chemical, or heat, the phosphate backbone at each abasic site can be cleaved.

15 The glycosylase utilized in the following procedures was uracil-DNA glycosylase (UDG). Uracil bases were incorporated into DNA fragments in each position that a thymine base would normally occupy by amplifying a DNA target sequence in the presence of uracil. Each uracil substituted DNA amplicon was incubated with UDG, which cleaved each uracil base
20 in the amplicon, and was then subjected to conditions that effected backbone cleavage at each abasic site, which produced DNA fragments. DNA fragments were subjected to MALDI-TOF mass spectrometry analysis. Genetic variability in the target DNA was then assessed by analyzing mass spectra.

25 Glycosylases specific for nucleotide analogs or modified nucleotides, as described herein, can be substituted for UDG in the following procedures. The glycosylase methods described hereafter, in conjunction with phosphate backbone cleavage and MALDI, can be used to analyze DNA fragments for the purposes of SNP scanning, bacteria

typing, methylation analysis, microsatellite analysis, genotyping, and nucleotide sequencing and re-sequencing.

A. Genotyping

- A glycosylase procedure was used to genotype the DNA sequence encoding UCP-2 (Uncoupling Protein 2). The sequence for UCP-2 is deposited in GenBank under accession number AF096289. The sequence variation genotyped in the following procedure was a cytosine (C-allele) to thymine (T-allele) variation at nucleotide position 4790, which results in a alanine to valine mutation at position 55 in the UCP-2 polypeptide.
- 5 DNA was amplified using a PCR procedure with a 50 µL reaction volume containing of 5 pmol biotinylated primer having the sequence 5'-TGCTTATCCCTGTAGCTACCCTGTCTGGCCTTGCAGATCCAA-3' (SEQ ID NO: 91), 15 pmol non-biotinylated primer having the sequence 5'-AGCGGATAACAATTCACACAGGCCATCACACCGCGGTACTG-3' (SEQ 10 ID NO: 92), 200 µM dATP, 200 µM dCTP, 200 µM dGTP, 600 µM dUTP (to fully replace dTTP), 1.5 mM to 3 mM MgCl₂, 1 U of HotStarTaq polymerase, and 25 ng of CEPH DNA. Amplification was effected with 15 45 cycles at an annealing temperature of 56°C.

- The amplification product was then immobilized onto a solid support by incubating 50 µL of the amplification reaction with 5 µL of prewashed Dynabeads for 20 minutes at room temperature. The supernatant was removed, and the beads were incubated with 50 µL of 0.1 M NaOH for 5 minutes at room temperature to denature the double-stranded PCR product in such a fashion that single-stranded DNA was 20 linked to the beads. The beads were then neutralized by three washes with 50 µL 10 mM TrisHCl (pH 8). The beads were resuspended in 10 µL of a 60mM TrisHCl/1mM EDTA (pH 7.9) solution, and 1 U uracil DNA glycosylase was added to the solution for 45 minutes at 37°C to remove 25 uracil nucleotides present in the single-stranded DNA linked to the beads.

The beads were then washed two times with 25 μ L of 10 mM TrisHCl (pH 8) and once with 10 μ L of water. The biotinylated strands were then eluted from the beads with 12 μ L of 2 M NH₄OH at 60°C for 10 minutes. The backbone of the DNA was cleaved by incubating the samples for 10 min at 95°C (with a closed lid), and ammonia was evaporated from the samples by incubating the samples for 11 min at 80°C.

5 The cleavage fragments were then analyzed by MALDI-TOF mass spectrometry as described in Example 4. The T-allele generated a unique fragment of 3254 Daltons. The C-allele generated a unique fragment of 10 4788 Daltons. These fragement were distinguishable in mass spectra. Thus, the above-identified procedure was successfully utilized to genotype individuals heterozygous for the C-allele and T-allele in UCP-2.

B. Glycosylase Analysis Utilizing Pooled DNA Samples

15 The glycosylase assay was conducted using pooled samples to detect genetic variability at the UCP-2 locus. DNA of known genotype was pooled from eleven individuals and was diluted to a fixed concentration of 5 ng/ μ L. The procedure provided in Example 3A was followed using 2 pmol of forward primer having a sequence of 5'-
CCCCAGTCACGACGTTGTAAAACGTCTGGCCTTGCAGATCCAAG- 3'
20 (SEQ ID NO: 93) and 15 pmol of reverse primer having the sequence 5'-
AGCGGATAACAATTCACACAGGCCATCACACCGCGGTACTG-3' (SEQ ID NO: 94). In addition, 5 pmol of biotinylated primer having the sequence 5'bioCCCAGTCACGACGTTGTAAAACG 3' (SEQ ID NO: 97) may be introduced to the PCR reaction after about two cycles. The
25 fragments were analyzed via MALDI-TOF mass spectroscopy (Example 4). As determined in Example 3A, the T-allele, which generated a unique fragment of 3254 Daltons, could be distinguished in mass spectra from the C-allele, which generated a unique fragment of 4788 Daltons. Allelic frequency in the pooled samples was quantified by integrating the area

under each signal corresponding to an allelic fragment. Integration was accomplished by hand calculations using equations well known to those skilled in the art. In the pool of eleven samples, this procedure suggested that 40.9% of the individuals harbored the T allele and 59.09% of the
5 individuals harbored the C allele.

C. Glycosylase-Mediated Microsatellite Analysis

A glycosylase procedure was utilized to identify microsatellites of the Bradykinin Receptor 2 (BKR-2) sequence. The sequence for BKR-2 is deposited in GenBank under accession number X86173. BKR-2 includes
10 a SNP in the promoter region, which is a C to T variation, as well as a SNP in a repeated unit, which is a G to T variation. The procedure provided in Example 3A was utilized to identify the SNP in the promotor region, the SNP in the microsatellite repeat region, and the number of repeated units in the microsatellite region of BKR-2. Specifically, a
15 forward PCR primer having the sequence 5'-
CTCCAGCTGGGCAGGAGTGC-3' (SEQ ID NO: 95) and a reverse primer having the sequence 5'-CACTTCAGTCGCTCCCT-3' (SEQ ID NO: 96)
were utilized to amplify BKR-2 DNA in the presence of uracil. The
amplicon was fragmented by UDG followed by backbone cleavage. The
20 cleavage fragments were analyzed by MALDI-TOF mass spectrometry as described in Example 4.

With regard to the SNP in the BKR-2 promotor region having a C to T variation, the C-allele generated a unique fragment having a mass of 7342.4 Daltons and the T-allele generated a unique fragment having a
25 mass of 7053.2 Daltons. These fragments were distinguishable in mass spectra. Thus, the above-identified procedure was successfully utilized to genotype individuals heterozygous for the C-allele and T-allele in the promotor region of BKR-2.

-77-

With regard to the SNP in the BKR-2 repeat region having a G to T variation, the T-allele generated a unique fragment having a mass of 1784 Daltons, which was readily detected in a mass spectrum. Hence, the presence of the T-allele was indicative of the G to T sequence variation in the repeat region of BKR-2.

In addition, the number of repeat regions was distinguished between individuals having two repeat sequences and individuals having three repeat sequences in BKR-2. The DNA of these individuals did not harbor the G to T sequence variation in the repeat sequence as each repeat sequence contained a G at the SNP locus. The number of repeat regions was determined in individual samples by calculating the area under a signal corresponding to a unique DNA fragment having a mass of 2771.6 Daltons. This signal in spectra generated from individuals having two repeat regions had an area that was thirty-three percent less than the area under the same signal in spectra generated from individuals having three repeat regions. Thus, the procedures discussed above can be utilized to genotype individuals for the number of repeat sequences present in BKR-2.

D. Bisulfite Treatment Coupled with Glycosylase Digestion

Bisulfite treatment of genomic DNA can be utilized to analyze positions of methylated cytosine residues within the DNA. Treating nucleic acids with bisulfite deaminates cytosine residues to uracil residues, while methylated cytosine remains unmodified. Thus, by comparing the sequence of a PCR product generated from genomic DNA that is not treated with bisulfite with the sequence of a PCR product generated from genomic DNA that is treated with bisulfite, the degree of methylation in a nucleic acid as well as the positions where cytosine is methylated can be deduced.

-78-

Genomic DNA (2 µg) was digested by incubation with 1 µL of a restriction enzyme at 37°C for 2 hours. An aliquot of 3 M NaOH was added to yield a final concentration of 0.3M NaOH in the digestion solution. The reaction was incubated at 37°C for 15 minutes followed by 5 treatment with 5.35M urea, 4.44M bisulfite, and 10mM hydroquinone, where the final concentration of hydroquinone is 0.5 mM.

The sample that was treated with bisulfite (sample A) was compared to the same digestion sample that had not undergone bisulfite treatment (sample B). After sample A was treated with bisulfite as 10 described above, sample A and sample B were amplified by a standard PCR procedure. The PCR procedure included the step of overlaying each sample with mineral oil and then subjecting the sample to thermocycling (20 cycles of 15 minutes at 55°C followed by 30 seconds at 95°C). The PCR reaction contained four nucleotide bases, C, A, G, and U. The 15 mineral oil was removed from each sample, and the PCR products were purified with glassmilk. Sodium iodide (3 volumes) and glassmilk (5 µL) were added to samples A and B. The samples were then placed on ice for 8 minutes, washed with 420 µL cold buffer, centrifuged for 10 seconds, and the supernatant fractions were removed. This process was 20 repeated twice and then 25 µL of water was added. Samples were incubated for 5 minutes at 37 °C, were centrifuged for 20 seconds, and the supernatant fraction was collected, and then this incubation/centrifugation/supernatant fraction collection procedure was repeated. 50 µL 0.1 M NaOH was then added to the samples to denature 25 the DNA. The samples were incubated at room temperature for 5 minutes, washed three times with 50 µL of 10 mM TrisHCl (pH 8), and resuspended in 10 µL 60mM TrisHCl/1mM EDTA, pH 7.9.

The sequence of PCR products from sample A and sample B were then treated with 2U of UDG (MBI Fermentas) and then subjected to

-79-

backbone cleavage, as described herein. The resulting fragments from each of sample A and sample B were analyzed by MALDI-TOF mass spectroscopy as described in Example 4. Sample A gave rise to a greater number of fragments than the number of fragments arising from sample
5 B, indicative that the nucleic acid harbored at least one methylated cytosine moiety.

EXAMPLE 7

Fen-Ligase-Mediated Haplotyping

Haplotyping procedures permit the selection of a fragment from one of an
10 individual's two homologous chromosomes and to genotype linked SNPs on that fragment. The direct resolution of haplotypes can yield increased information content, improving the diagnosis of any linked disease genes or identifying linkages associated with those diseases. In previous studies, haplotypes were typically reconstructed indirectly through
15 pedigree analysis (in cases where pedigrees were available) through laborious and unreliable allele-specific PCR or through single-molecule dilution methods well known in the art.

A haplotyping procedure was used to determine the presence of two SNPs, referred to as SNP1 and SNP2, located on one strand in a DNA
20 sample. The haplotyping procedure used in this assay utilized Fen-1, a site-specific "flap" endonuclease that cleaves DNA "flaps" created by the overlap of two oligonucleotides hybridized to a target DNA strand. The two overlapping oligonucleotides in this example were short arm and long arm allele-specific adaptors. The target DNA was an amplified nucleic
25 acid that had been denatured and contained SNP1 and SNP2.

The short arm adaptor included a unique sequence not found in the target DNA. The 3' distal nucleotide of the short arm adaptor was identical to one of the SNP1 alleles. Moreover, the long arm adaptor included two regions: a 3' region complementary to the short arm and a

-80-

5'gene-specific region complementary to the fragment of interest adjacent to the SNP. If there was a match between the adaptor and one of the homologues, the Fen enzyme recognized and cleaved the overlapping flap. The short arm of the adaptor was then ligated to the remainder of 5 the target fragment (minus the SNP site). This ligated fragment was used as the forward primer for a second PCR reaction in which only the ligated homologue was amplified. The second PCR product (PCR2) was then analyzed by mass spectrometry. If there was no match between the adaptors and the target DNA, there was no overlap, no cleavage by Fen- 10 1, and thus no PCR2 product of interest.

If there was more than one SNP in the sequence of interest, the second SNP (SNP2) was found by using an adaptor that was specific for SNP2 and hybridizing the adaptor to the PCR2 product containing the first SNP. The Fen-ligase and amplification procedures were repeated for the 15 PCR2 product containing the first SNP. If the amplified product yielded a second SNP, then SNP1 and SNP2 were on the same fragment.

If the SNP is unknown, then four allele-specific adaptors (e.g. C, G, A, and T) can be used to hybridize with the target DNA. The substrates are then treated with the Fen-ligase protocol, including amplification. The 20 PCR2 products may be analyzed by PROBE, as described herein, to determine which adaptors were hybridized to the DNA target and thus identify the SNPs in the sequence.

A Fen-ligase assay was used to detect two SNPs present in Factor VII. These SNPs are located 814 base pairs apart from each other. SNP1 25 was located at position 8401 (C to T), and SNP2 was located at 9215 (G to A) (SEQ ID #).

A. First Amplification Step

A PCR product (PCR1) was generated for a known heterozygous individual at SNP1, a short distance from the 5' end of the SNP.

-81-

- Specifically, a 10 μ L PCR reaction was performed by mixing 1.5 mM MgCl₂, 200 μ M of each dNTP, 0.5 U HotStar polymerase, 0.1 μ M of a forward primer having the sequence 5'-GCG CTC CTG TCG GTG CCA (SEQ ID NO: 56), 0.1 μ M of a reverse primer having the sequence 5'-GCC 5 TGA CTG GTG GGG CCC (SEQ ID NO: 57), and 1 ng of genomic DNA. The annealing temperature was 58°C, and the amplification process yielded fragments that were 861 bp in length.

- The PCR1 reaction mixture was divided in half and was treated with an exonuclease 1/SAP mixture (0.22 μ L mixture/5 μ L PCR1 reaction) 10 which contained 1.0 μ L SAP and 0.1 μ L exon1. The exonuclease treatment was done for 30 minutes at 37°C and then 20 minutes at 85°C to denature the DNA.

B. Adaptor Oligonucleotides

- A solution of allele-specific adaptors (C and T), containing of one 15 long and one short oligonucleotide per adaptor, was prepared. The long arm and short arm oligonucleotides of each adaptor (10 μ M) were mixed in a 1:1 ratio and heated for 30 seconds at 95°C. The temperature was reduced in 2°C increments to 37°C for annealing. The C-adaptor had a short arm sequence of 5'-CAT GCA TGC ACG GTC (SEQ ID NO: 58) and 20 a long arm sequence of 5'-CAG AGA GTA CCC CTC GAC CGT GCA TGC ATG (SEQ ID NO: 59). Hence, the long arm of the adaptor was 30 bp (15 bp gene-specific), and the short arm was 15bp. The T-adaptor had a short arm sequence of 5'-CAT GCA TGC ACG GTT (SEQ ID NO: 60) and a long arm sequence of 5'-GTA CGT ACG TGC CAA CTC CCC ATG AGA 25 GAC (SEQ ID NO: 61). The adaptor could also have a hairpin structure in which the short and long arm are separated by a loop containing of 3 to 10 nucleotides (SEQ ID NO: 118).

C. FEN-ligase reaction

-82-

In two tubes (one tube for each allele-specific adaptor per sample) was placed a solution (Solution A) containing of 3.5 μ l 10 mM 16%PEG/50 mM MOPS, 1.2 μ l 25 mM MgCl₂, 1.5 μ l 10X Ampligase Buffer, and 2.5 μ l PCR1. Each tube containing Solution A was incubated 5 at 95°C for 5 minutes to denature the PCR1 product. A second solution (Solution B) containing of 1.65 μ l Ampligase (Thermostable ligase, Epicentre Technologies), 1.65 μ l 200ng/ μ l MFEN (from *Methanococcus jannaschii*), and 3.0 μ l of an allele specific adaptor (C or T) was prepared. Thus, different variations of Solution B, each variation containing of 10 different allele-specific adaptors, were made. Solution B was added to Solution A at 95°C and incubated at 55°C for 3 hours. The total reaction volume was 15.0 μ l per adaptor-specific reaction. For a bi-allelic system, 2 x 15.0 μ l reactions were required.

The Fen-ligase reaction in each tube was then deactivated by 15 adding 8.0 μ l 10 mM EDTA. Then, 1.0 μ l exoIII/Buffer (70%/30%) solution was added to each sample and incubated 30 minutes at 37°C, 20 minutes at 70°C (to deactivate exoIII), and 5 minutes at 95°C (to denature the sample and dissociate unused adaptor from template). The samples were cooled in an ice slurry and purified on UltraClean PCR 20 Clean-up (MoBio) spin columns which removed all fragments less than 100 base pairs in length. The fragments were eluted with 50 μ l H₂O.

D. Second Amplification Step

A second amplification reaction (PCR2) was conducted in each sample tube using the short arm adaptor (C or T) sequence as the forward 25 primer (minus the SNP1 site). Only the ligated homologue was amplified. A standard PCR reaction was conducted with a total volume of 10.0 μ l containing of 1X Buffer (final concentration), 1.5 mM final concentration MgCl₂, 200 μ M final concentration dNTPs, 0.5 U HotStar polymerase, 0.1 μ M final concentration forward primer 5'-CAT GCA TGC ACG GT (SEQ ID

-83-

NO: 62), 0.1 μ M final concentration reverse primer 5'-GCC TGA CTG GTG GGG CCC (SEQ ID NO: 63), and 1.0 μ l of the purified FEN-ligase reaction solution. The annealing temperature was 58°C. The PCR2 product was analyzed by MALDI TOF mass spectroscopy as described in Example 4.

- 5 The mass spectrum of Fen SNP1 showed a mass of 6084.08 Daltons, representing the C allele.

E. Genotyping Additional SNPs

The second SNP (SNP2) can be found by using an adaptor that is specific for SNP2 and hybridizing that adaptor to the PCR2 product

- 10 containing the first SNP. The Fen-ligase and amplification procedures are repeated for the PCR2 product containing the first SNP. If the amplified product yields a second SNP, then SN1 and SN2 are on the same fragment. The mass spectrum of SNP2, representing the T allele, showed a mass of 6359.88 Daltons.

- 15 This assay can also be performed upon pooled DNA to yield haplotype frequencies as described herein. The Fen-ligase assay can be used to analyze multiplexes as described herein.

EXAMPLE 8

Nickase-Mediated Sequence Analysis

- 20 A DNA nickase, or DNase, was used to recognize and cleave one strand of a DNA duplex. Two nickases used were NY2A nickase and NYS1 nickase (Megabase) which cleave DNA at the following sites:

NY2A: 5'...R AG...3'

3'...Y \downarrow TC...5' where R = A or G and Y = C or T

- 25 NYS1: 5'... \downarrow CC[A/G/T]...3'

3'... GG[T/C/A]...5'.

-84-

A. Nickase Digestion

Tris-HCl (10 mM), KCl (10 mM, pH 8.3), magnesium acetate (25 mM), BSA (1 mg/mL), and 6 U of Cvi NY2A or Cvi NYS1 Nickase (Megabase Research) were added to 25 pmol of double-stranded
5 oligonucleotide template having a sequence of 5'-CGC AGG GTT TCC
TCG TCG CAC TGG GCA TGT G-3' (SEQ ID NO: 90, Operon, Alameda,
CA) synthesized using standard phosphoramidite chemistry . With a total
volume of 20 μ L, the reaction mixture was incubated at 37°C for 5 hours,
and the digestion products were purified using ZipTips (Millipore, Bedford,
10 MA) as described in Example 5. The samples were analyzed by MALTY-
TOM mass spectroscopy as described in Example 1. The nickase Cvi
NY2A yielded three fragments with masses 4049.76 Daltons, 5473.14
Daltons, and 9540.71 Daltons. The Cvi NYS1 nickase yielded fragments
with masses 2063.18 Daltons, 3056.48 Daltons, 6492.81 Daltons, and
15 7450.14 Daltons.

B. Nickase Digestion of Pooled Samples

DQA (HLA ClassII-DQ Alpha, expected fragment size = 225bp) was amplified from the genomic DNA of 100 healthy individuals. DQA was amplified using standard PCR chemistry in a reaction having a total
20 volume of 50 μ L containing of 10 mM Tris-HCl, 10 mM KCl (pH 8.3), 2.5
mM MgCl₂, 200 μ M of each dNTP, 10 pmol of a forward primer having
the sequence 5'-GTG CTG CAG GTG TAA ACT TGT ACC AG-3'(SEQ ID
NO: 64), 10 pmol of a reverse primer having the sequence 5'-CAC GGA
TCC GGT AGC AGC GGT AGA GTT G-3'(SEQ ID NO: 65), 1 U DNA
25 polymerase (Stoffel fragment, Perkin Elmer), and 200ng human genomic
DNA (2ng DNA/individual). The template was denatured at 94°C for 5
minutes. Thermal cycling was continued with a touch-down program that
included 45 cycles of 20 seconds at 94°C, 30 seconds at 56°C, 1

-85-

minute at 72°C, and a final extension of 3 minutes at 72°C. The crude PCR product was used in the subsequent nickase reaction.

- The unpurified PCR product was subjected to nickase digestion. Tris-HCl (10 mM), KCl (10 mM, pH 8.3), magnesium acetate (25mM), 5 BSA (1 mg/mL), and 5 U of Cvi NY2A or Cvi NYS1 Nickase (Megabase Research) were added to 25 pmol of the amplified template with a total reaction volume of 20 μ L. The mixture was then incubated at 37°C for 5 hours. The digestion products were purified with either ZipTips (Millipore, Bedford, MA) as described in Example 5. The samples were analyzed by 10 MALDI-TOF mass spectroscopy as described in Example 4. This assay can also be used to do multiplexing and standardless genotyping as described herein.

- To simplify the nickase mass spectrum, the two complementary strands can be separated after digestion by using a single-stranded 15 undigested PCR product as a capture probe. This probe (preparation shown below in Example 8C) can be hybridized to the nickase fragments in hybridization buffer containing 200 mM sodium citrate and 1% blocking reagent (Boehringer Mannheim). The reaction is heated to 95°C for 5 minutes and cooled to room temperature over 30 minutes by using a 20 thermal cycler (PTC-200 DNA engine, MJ Research, Waltham, MA). The capture probe-nickase fragment is immobilized on 140 μ g of streptavidin-coated magnetic beads. The beads are subsequently washed three times with 70 mM ammonium citrate. The captured single-stranded nickase 25 fragments are eluted by heating to 80°C for 5 minutes in 5 μ L of 50 mM ammonium hydroxide.

C. Preparation of Capture Probe

The capture probe is prepared by amplifying the human β -globin gene (3' end of intron 1 to 5' end of exon 2) via PCR methods in a total volume of 50 μ L containing of GeneAmp 1XPCR Buffer II, 10 mM Tris-

-86-

- HCl, pH 8.3, 50 mM KCl, 2 mM MgCl₂, 0.2 mM dNTP mix, 10pmol of each primer (forward primer 5'-ACTGGGCATGTGGAGACAG-3' (SEQ ID NO: 66) and biotinylated reverse primer bio5'-GCACTTTCTTGCCATGAG-3' (SEQ ID: 67), 2 U of AmpliTaq Gold, and 200 ng of human genomic DNA. The template is denatured at 94°C for 8 minutes. Thermal cycling is continued with a touch-down program that included 11 cycles of 20 seconds at 94°C, 30 seconds at 64°C, 1 minute at 72°C; and a final extension of 5 minutes at 72°C. The amplicon is purified using UltraClean™ PCR clean-up kit (MO Bio Laboratories, Solano Beach, CA).

10

EXAMPLE 9

Multiplex Type IIS SNP Assay

- A Type IIS assay was used to identify human gene sequences with known SNPs. The Type IIS enzyme used in this assay was Fok I which effected double-stranded cleavage of the target DNA. The assay involved the steps of amplification and Fok I treatment of the amplicon. In the amplification step, the primers were designed so that each PCR product of a designated gene target was less than 100 bases such that a Fok I recognition sequence was incorporated at the 5' and 3' end of the amplicon. Therefore, the fragments that were cleaved by Fok I included a center fragment containing the SNP of interest.

- Ten human gene targets with known SNPs were analyzed by this assay. Sequences of the ten gene targets, as well as the primers used to amplify the target regions, are found in Table 5. The ten targets were lipoprotein lipase, prothrombin, factor V, cholesterol ester transfer protein (CETP), factor VII, factor XIII, HLA-H exon 2, HLA-H exon 4, methylenetetrahydrofolate reductase (MTHR), and P53 exon 4 codon 72.

Amplification of the ten human gene sequences were carried out in a single 50 µL volume PCR reaction with 20 ng of human genomic DNA

-87-

template in 5 PCR reaction tubes. Each reaction vial contained 1X PCR buffer (Qiagen), 200 μ M dNTPs, 1U Hotstar Taq polymerase (Qiagen), 4 mM MgCl₂, and 10pmol of each primer. US8, having sequence of 5'TCAGTCACGACGTT3' (SEQ ID NO: 68), and US9, having sequence of 5'CGGATAACAATTTC3' (SEQ ID NO: 69), were used for the forward and reverse primers respectively. Moreover, the primers were designed such that a Fok I recognition site was incorporated at the 5' and 3' ends of the amplicon. Thermal cycling was performed in 0.2 mL tubes or a 96 well plate using a MJ Research Thermal Cycler (calculated temperature) with the following cycling parameters: 94°C for 5 minutes; 45 cycles: 94°C for 20 seconds, 56°C for 20 seconds, 72°C for 60 seconds; and 72°C for 3 minutes.

Following PCR, the sample was treated with 0.2 U Exonuclease I (Amersham Pharmacia) and S Alkaline Phosphotase (Amersham Pharmacia) to remove the unincorporated primers and dNTPs. Typically, 0.2 U of exonuclease I and SAP were added to 5 μ L of the PCR sample. The sample was then incubated at 37°C for 15 minutes. Exonuclease I and SAP were then inactivated by heating the sample up to 85°C for 15 minutes. Fok I digestion was performed by adding 2 U of Fok I (New England Biolab) to the 5 μ L PCR sample and incubating at 37°C for 30 minutes. Since the Fok I restriction sites are located on both sides of the amplicon, the 5' and 3' cutoff fragments have higher masses than the center fragment containing the SNP. The sample was then purified by anion exchange and analyzed by MALDI-TOF mass spectrometry as described in Example 4. The masses of the gene fragments from this multiplexing experiment are listed in Table 6. These gene fragments were resolved in mass spectra thereby allowing multiplex analysis of sequence variability in these genes.

30 Table 5
 Genes for Multiplex Type IIS Assay

-88-

	Gene	Sequence	Seq. ID No.	Primers	Seq. ID No.
5	Lipoprotein Lipase (Asn291Ser)	cctttgagaa <u>aggcgtctgc ttgatgttta</u> <u>gaaaagaaccg ctgcacaaat</u> <u>ctggctatg agatca[a>g]taa agtcagagcc</u> <u>aaaagaagca gaaaaatgtt</u>	98-99	5' caatttcacacacggcgatgcitct 5' caatttcacacacggcgatgcitct 5' caatttcacacacggcgatgcitct 5' caatttcacacacggcgatgcitct	70 71
	Prothrombin	26731 <u>gaatttttttgtgtttttaaaactatgggt</u> <u>tcccaataaa agtqacitc</u> 26781 <u>atgc[g>a]gcctc aatqctcccc</u> <u>gtgtatcca tggccaccc tcgtggctca</u>	100-101	5' tcagtacacgcgttggatgccaa 5' taaaatgtactctcagc 3' 5' cggataacaatttcggatgcact 5' gggagcattgaggc 3'	72 73
	Factor V (Arg506Gln)	taataggact <u>acctttaatc tgtaagagca</u> <u>gatcccttggaa caggc(g>a)aggaa</u> <u>atacaggat tttgtccttg aagtaaacctt tcag</u>	102-103	5' tcagtacacgcgttggatgagca 5' gatcccttggacaggc 3' 5' cggataacaatttcggatggaca 5' aaataccctgtattcc 3'	74 75
10	Cholesterol ester transfer protein (CETP) (I405V)	1261 ctcacatgg gcatttggat <u>gcagagcage</u> <u>tcgcgatcc[g>a] tccagagctt</u> 1311 <u>cctgcgtca atqatcaccc ctgtgggcatt</u> <u>ccctgaggtc atgtctgtt</u>	104-105	5' tcagtacacgcgttggatgcaga 5' gcaatcccgatgc 3' 5' cagcggtgtatggatgcagg 5' asgcctctgg 3'	76 77
	Factor VII (R353Q)	1221 <u>agcaaggact cctgcaaggg ggacatgttga</u> <u>ggcccccacatg ccacccacta</u> 1271 <u>cc[a>g]gggcacq tggatctqa</u> <u>cgggcategt cagctggggc cagggtctgcg</u>	106-107	5' tcagtacacgcgttggatgcaca 5' catggccacccaciac 3' 5' cggataacaatttcggatgcagg 5' tcgggttccacac 3'	78 79
	Factor XIII (V34L)	111 <u>caataactct aatgcacggg aagatgaccct</u> <u>ggcccacatq gqacttcagg</u> 161 <u>gc[g>t]tggtgcc cgggggcgtc</u> <u>sacctgcaag qtatqagcat acccccccttc</u>	108-109	5' tcagtacacgcgttggatgcaca 5' cagttggatctc 3' 5' gctcataaccttgcaggatgacg 5' 3'	80 81
15	HLA-H exon 2 (His63Asp)	361 <u>ttgaagctttggctecgtg gatgaccagg</u> <u>tggtcgatgttctatgtt[c>g]at</u> 411 <u>gagagtgcgcq tggtggagcc ccgaactccaa</u> <u>tgggtttccaa gtataatttc</u>	110-111	5' tcagtacacgcgttggatgacca 5' gcttttcgtgttc 3' 5' tacatggatgttggggatgcaca 5' cggcqacttc 3'	82 83
	HLA-H exon 4 (Cys282Tyr)	1021 <u>ggataacctt ggctgtaccc cctggggggag</u> <u>agcagatata <u>tacgt[g>a]ccag</u></u> 1071 <u>gtggggccacc caggcccttggaa tcagccccctc</u> <u>attgtgtatct gggagccctc</u>	112-113	5' tcagtacacgcgttggatgggg 5' agacgagatataacgt 5' ggggggtgtatccaaaggatggat 5' gctccac 3'	84 85

-89-

Gene	Sequence	Seq. ID No.	Primers	Seq. ID No.
Methylenetetrahydrofolatereductase (MTHFR) (Ala222Val)	761 <u>ttaaacactt</u> <u>ggggaa</u> <u>ggaa</u> <u>gtatctccaa</u> gag(c>)cgattt catcatcacg 811 <u>cacgttttc</u> <u>ttagggctga</u> <u>cacattttc</u>	114-115	5' tcagtacgcgcgttggatgggg agaccaaaaaataacgt 3' 5' gggggggctgatccaggatgggt gctccac 3'	86 87
P53 Exon4 Codon 72 (Arg72Pro)	12101 tccagatqaa <u>gctccca</u> <u>gg</u> tgccagggc tgctcccc(g>c)c gtggcccttg 12151 <u>caccac</u> cagec <u>tcctacacc</u> ggggccctcg	116-117	5' gatggaaatcccaggatgccag aggc 3' 5' ggggggggtttaggatgcgtcg gtac 3'	88 89

-89/a-

Table 6
The mass of Center Fragments for Ten Different SNP Typing by
IIS Assay

Gene	LPL(^A 291 ^G)	Prothrombin	FV(^A 506 ^G)	CETP(^A 405 ^G)	FVIII(^A 353 ^G)	FXIII(^A 34)
Genotype	A	G	A	G	A	G
+ strand mass (Da)	6213	6229	5845	5829	5677	5661
- strand mass (Da)	6129	6114	5949	5964	5472	5487

Gene	HhaII	HhaII	MTHFR(^A 222 ^G)	P53exon4(^A 97 ^G)
Genotype	C	G	A	C
+ strand mass (Da)	5889	5929	4392	4376
- strand mass (Da)	5836	5796	4319	4334

-90-

EXAMPLE 10

Exemplary use of parental medical history parameter for stratification of healthy database

A healthy database can be used to associate a disease state with a specific allele (SNP) that has been found to show a strong association between age and the allele, in particular the homozygous genotype. The method involves using the same healthy database used to identify the age dependent association, however stratification is by information given by the donors about common disorders from which their parents suffered (the donor's familial history of disease). There are three possible answers a donor could give about the health status of their parents: neither were affected, one was affected or both were affected. Only donors above a certain minimum age, depending on the disease, are utilized, as the donors parents must be old enough to have exhibited clinical disease phenotypes. The genotype frequency in each of these groups is determined and compared with each other. If there is an association of the marker in the donor to a disease the frequency of the heterozygous genotype will be increased. The frequency of the homozygous genotype should not increase, as it should be significantly underrepresented in the healthy population.

-91-

EXAMPLE 11

Method and Device for Identifying a Biological Sample

Description

In accordance with the present invention, a method and device for identifying a biological sample is provided. Referring now to FIG. 24, an apparatus 10 for identifying a biological sample is disclosed. The apparatus 10 for identifying a biological sample generally comprises a mass spectrometer 15 communicating with a computing device 20. In a preferred embodiment, the mass spectrometer may be a MALDI-TOF mass spectrometer manufactured by Bruker-Franzen Analytik GmbH; however, it will be appreciated that other mass spectrometers can be substituted. The computing device 20 is preferably a general purpose computing device. However, it will be appreciated that the computing device could be alternatively configured, for example, it may be integrated with the mass spectrometer or could be part of a computer in a larger network system.

The apparatus 10 for identifying a biological sample may operate as an automated identification system having a robot 25 with a robotic arm 27 configured to deliver a sample plate 29 into a receiving area 31 of the mass spectrometer 15. In such a manner, the sample to be identified may be placed on the plate 29 and automatically received into the mass spectrometer 15. The biological sample is then processed in the mass spectrometer to generate data indicative of the mass of DNA fragments in the biological sample. This data may be sent directly to computing device 20, or may have some preprocessing or filtering performed within the mass spectrometer. In a preferred embodiment, the mass spectrometer 15 transmits unprocessed and unfiltered mass spectrometry data to the computing device 20. However, it will be appreciated that the analysis in the computing device may be adjusted to accommodate preprocessing or filtering performed within the mass spectrometer.

Referring now to FIG. 25, a general method 35 for identifying a biological sample is shown. In method 35, data is received into a computing device from a test instrument in block 40. Preferably the data is received in a raw, unprocessed and unfiltered form, but alternatively may have some form of

filtering or processing applied. The test instrument of a preferred embodiment is a mass spectrometer as described above. However, it will be appreciated that other test instruments could be substituted for the mass spectrometer.

The data generated by the test instrument, and in particular the mass spectrometer, includes information indicative of the identification of the biological sample. More specifically, the data is indicative of the DNA composition of the biological sample. Typically, mass spectrometry data gathered from DNA samples obtained from DNA amplification techniques are noisier than, for example, those from typical protein samples. This is due in part because protein samples are more readily prepared in more abundance, and protein samples are more easily ionizable as compared to DNA samples. Accordingly, conventional mass spectrometer data analysis techniques are generally ineffective for DNA analysis of a biological sample. To improve the analysis capability so that DNA composition data can be more readily discerned, a preferred embodiment uses wavelet technology for analyzing the DNA mass spectrometry data. Wavelets are an analytical tool for signal processing, numerical analysis, and mathematical modeling. Wavelet technology provides a basic expansion function which is applied to a data set. Using wavelet decomposition, the data set can be simultaneously analyzed in the time and frequency domains. Wavelet transformation is the technique of choice in the analysis of data that exhibit complicated time (mass) and frequency domain information, such as MALDI-TOF DNA data. Wavelet transforms as described herein have superior denoising properties as compared to conventional Fourier analysis techniques. Wavelet transformation has proven to be particularly effective in interpreting the inherently noisy MALDI-TOF spectra of DNA samples. In using wavelets, a "small wave" or "scaling function" is used to transform a data set into stages, with each stage representing a frequency component in the data set. Using wavelet transformation, mass spectrometry data can be processed, filtered, and analyzed with sufficient discrimination to be useful for identification of the DNA composition for a biological sample.

Referring again to FIG. 25, the data received in block 40 is denoised in block 45. The denoised data then has a baseline correction applied in block 50.

-93-

A baseline correction is generally necessary as data coming from the test instrument, in particular a mass spectrometer instrument, has data arranged in a generally exponentially decaying manner. This generally exponential decaying arrangement is not due to the composition of the biological sample, but is a
5 result of the physical properties and characteristics of the test instrument, and other chemicals involved in DNA sample preparation. Accordingly, baseline correction substantially corrects the data to remove a component of the data attributable to the test system, and sample preparation characteristics.

After denoising in block 45 and the baseline correction in block 50, a
10 signal remains which is generally indicative of the composition of the biological sample. However, due to the extraordinary discrimination required for analyzing the DNA composition of the biological sample, the composition is not readily apparent from the denoised and corrected signal. For example, although the signal may include peak areas, it is not yet clear whether these "putative" peaks
15 actually represent a DNA composition, or whether the putative peaks are result of a systemic or chemical aberration. Further, any call of the composition of the biological sample would have a probability of error which would be unacceptable for clinical or therapeutic purposes. In such critical situations, there needs to be a high degree of certainty that any call or identification of the sample is
20 accurate. Therefore, additional data processing and interpretation is necessary before the sample can be accurately and confidently identified.

Since the quantity of data resulting from each mass spectrometry test is typically thousands of data points, and an automated system may be set to perform hundreds or even thousands of tests per hour, the quantity of mass
25 spectrometry data generated is enormous. To facilitate efficient transmission and storage of the mass spectrometry data, block 55 shows that the denoised and baseline corrected data is compressed.

In a preferred embodiment, the biological sample is selected and processed to have only a limited range of possible compositions. Accordingly, it
30 is therefore known where peaks indicating composition should be located, if present. Taking advantage of knowing the location of these expected peaks, in block 60 the method 35 matches putative peaks in the processed signal to the

-94-

location of the expected peaks. In such a manner, the probability of each putative peak in the data being an actual peak indicative of the composition of the biological sample can be determined. Once the probability of each peak is determined in block 60, then in block 65 the method 35 statistically determines 5 the composition of the biological sample, and determines if confidence is high enough to calling a genotype.

Referring again to block 40, data is received from the test instrument, which is preferably a mass spectrometer. In a specific illustration, FIG. 26 shows an example of data from a mass spectrometer. The mass spectrometer 10 data 70 generally comprises data points distributed along an x-axis 71 and a y-axis 72. The x-axis 71 represents the mass of particles detected, while the y-axis 72 represents a numerical concentration of the particles. As can be seen in FIG. 26, the mass spectrometry data 70 is generally exponentially decaying with data at the left end of the x-axis 73 generally decaying in an exponential manner 15 toward data at the heavier end 74 of the x-axis 71. However, the general exponential presentation of the data is not indicative of the composition of the biological sample, but is more reflective of systematic error and characteristics. Further, as described above and illustrated in FIG. 26, considerable noise exists in the mass spectrometry DNA data 70.

20 Referring again to block 45, where the raw data received in block 40 is denoised, the denoising process will be described in more detail. As illustrated in FIG. 25, the denoising process generally entails 1) performing a wavelet transformation on the raw data to decompose the raw data into wavelet stage coefficients; 2) generating a noise profile from the highest stage of wavelet 25 coefficients; and 3) applying a scaled noise profile to other stages in the wavelet transformation. Each step of the denoising process is further described below.

Referring now to FIG. 27, the wavelet transformation of the raw mass spectrometry data is generally diagramed. Using wavelet transformation techniques, the mass spectrometry data 70 is sequentially transformed into 30 stages. In each stage the data is represented in a high stage and a low stage, with the low stage acting as the input to the next sequential stage. For example, the mass spectrometry data 70 is transformed into stage 0 high data

-95-

- 82 and stage 0 low data 83. The stage 0 low data 83 is then used as an input to the next level transformation to generate stage 1 high data 84 and stage 1 low data 85. In a similar manner, the stage 1 low data 85 is used as an input to be transformed into stage 2 high data 86 and stage 2 low data 87. The
5 transformation is continued until no more useful information can be derived by further wavelet transformation. For example, in the preferred embodiment a 24-point wavelet is used. More particularly a wavelet commonly referred to as the Daubechies 24 is used to decompose the raw data. However, it will be appreciated that other wavelets can be used for the wavelet transformation.
10 Since each stage in a wavelet transformation has one-half the data points of the previous stage, the wavelet transformation can be continued until the stage n low data 89 has around 50 points. Accordingly, the stage n high 88 would contain about 100 data points. Since the preferred wavelet is 24 points long, little data or information can be derived by continuing the wavelet transformation
15 on a data set of around 50 points.

FIG. 28 shows an example of stage 0 high data 95. Since stage 0 high data 95 is generally indicative of the highest frequencies in the mass spectrometry data, stage 0 high data 95 will closely relate to the quantity of high frequency noise in the mass spectrometry data. In FIG. 29, an exponential
20 fitting formula has been applied to the stage 0 high data 95 to generate a stage 0 noise profile 97. In particular, the exponential fitting formula is in the format $A_0 + A_1 \text{EXP}(-A_2 m)$. It will be appreciated that other exponential fitting formulas or other types of curve fits may be used.

Referring now to FIG. 30, noise profiles for the other high stages are
25 determined. Since the later data points in each stage will likely be representative of the level of noise in each stage, only the later data points in each stage are used to generate a standard deviation figure that is representative of the noise content in that particular stage. More particularly, in generating the noise profile for each remaining stage, only the last five percent of the data points in each
30 stage are analyzed to determine a standard deviation number. It will be appreciated that other numbers of points, or alternative methods could be used to generate such a standard deviation figure.

-96-

The standard deviation number for each stage is used with the stage 0 noise profile (the exponential curve) 97 to generate a scaled noise profile for each stage. For example, FIG. 30 shows that stage 1 high data 98 has stage 1 high data 103 with the last five percent of the data points represented by area 99. The points in area 99 are evaluated to determine a standard deviation number indicative of the noise content in stage 1 high data 103. The standard deviation number is then used with the stage 0 noise profile 97 to generate a stage 1 noise profile.

In a similar manner, stage 2 high 100 has stage 2 high data 104 with the last five percent of points represented by area 101. The data points in area 101 are then used to calculate a standard deviation number which is then used to scale the stage 0 noise profile 97 to generate a noise profile for stage 2 data. This same process is continued for each of the stage high data as shown by the stage n high 105. For stage n high 105, stage n high data 108 has the last five percent of data points indicated in area 106. The data points in area 106 are used to determine a standard deviation number for stage n. The stage n standard deviation number is then used with the stage 0 noise profile 97 to generate a noise profile for stage n. Accordingly, each of the high data stages has a noise profile.

FIG. 31 shows how the noise profile is applied to the data in each stage. Generally, the noise profile is used to generate a threshold which is applied to the data in each stage. Since the noise profile is already scaled to adjust for the noise content of each stage, calculating a threshold permits further adjustment to tune the quantity of noise removed. Wavelet coefficients below the threshold are ignored while those above the threshold are retained. Accordingly, the remaining data has a substantial portion of the noise content removed.

Due to the characteristics of wavelet transformation, the lower stages, such as stage 0 and 1, will have more noise content than the later stages such as stage 2 or stage n. Indeed, stage n low data is likely to have little noise at all. Therefore, in a preferred embodiment the noise profiles are applied more aggressively in the lower stages and less aggressively in the later stages. For example, FIG. 31 shows that stage 0 high threshold is determined by multiplying

the stage 0 noise profile by a factor of four. In such a manner, significant numbers of data points in stage 0 high data 95 will be below the threshold and therefore eliminated. Stage 1 high threshold 112 is set at two times the noise profile for the stage 1 high data, and stage 2 high threshold 114 is set equal to

5 the noise profile for stage 2 high. Following this geometric progression, stage n high threshold 116 is therefore determined by scaling the noise profile for each respective stage n high by a factor equal to $(1/2^{n-2})$. It will be appreciated that other factors may be applied to scale the noise profile for each stage. For example, the noise profile may be scaled more or less aggressively to

10 accommodate specific systemic characteristics or sample compositions. As indicated above, stage n low data does not have a noise profile applied as stage n low data 118 is assumed to have little or no noise content. After the scaled noise profiles have been applied to each high data stage, the mass spectrometry data 70 has been denoised and is ready for further processing. A wavelet

15 transformation of the denoised signal results in the sparse data set 120 as shown in FIG. 31.

Referring again to FIG. 25, the mass spectrometry data received in block 40 has been denoised in block 45 and is now passed to block 50 for baseline correction. Before performing baseline correction, the artifacts introduced by the

20 wavelet transformation procedure are preferably removed. Wavelet transformation results vary slightly depending upon which point of the wavelet is used as a starting point. For example, the preferred embodiment uses the 24-point Daubechies-24 wavelet. By starting the transformation at the 0 point of the wavelet, a slightly different result will be obtained than if starting at points 1

25 or 2 of the wavelet. Therefore, the denoised data is transformed using every available possible starting point, with the results averaged to determine a final denoised and shifted signal. For example, FIG. 33 shows that the wavelet coefficient is applied 24 different times and then the results averaged to generate the final data set. It will be appreciated that other techniques may be

30 used to accommodate the slight error introduced due to wavelet shifting.

The formula 125 is generally indicated in FIG. 33. Once the signal has been denoised and shifted, a denoised and shifted signal 130 is generated as

-98-

shown in FIG. 58. FIG. 34 shows an example of the wavelet coefficient 135 data set from the denoised and shifted signal 130.

FIG. 36 shows that putative peak areas 145, 147, and 149 are located in the denoised and shifted signal 150. The putative peak areas are systematically 5 identified by taking a moving average along the signal 150 and identifying sections of the signal 150 which exceed a threshold related to the moving average. It will be appreciated that other methods can be used to identify putative peak areas in the signal 150.

Putative peak areas 145, 147 and 149 are removed from the signal 150 10 to create a peak-free signal 155 as shown in FIG. 37. The peak-free signal 155 is further analyzed to identify remaining minimum values 157, and the remaining minimum values 157 are connected to generate the peak-free signal 155.

FIG. 38 shows a process of using the peak-free signal 155 to generate a baseline 170 as shown in FIG. 39. As shown in block 162, a wavelet 15 transformation is performed on the peak-free signal 155. All the stages from the wavelet transformation are eliminated in block 164 except for the n low stage. The n low stage will generally indicate the lowest frequency component of the peak-free signal 155 and therefore will generally indicate the system exponential characteristics. Block 166 shows that a signal is reconstructed from the n low 20 coefficients and the baseline signal 170 is generated in block 168.

FIG. 39 shows a denoised and shifted data signal 172 positioned adjacent a correction baseline 170. The baseline correction 170 is subtracted from the denoised and shifted signal 172 to generate a signal 175 having a baseline 25 correction applied as shown in FIG. 40. Although such a denoised, shifted, and corrected signal is sufficient for most identification purposes, the putative peaks in signal 175 are not identifiable with sufficient accuracy or confidence to call the DNA composition of a biological sample.

Referring again to FIG. 25, the data from the baseline correction 50 is now compressed in block 55, the compression technique used in a preferred 30 embodiment is detailed in FIG. 41. In FIG. 41 the data in the baseline corrected data is presented in an array format 182 with x-axis points 183 having an associated data value 184. The x-axis is indexed by the non-zero wavelet

-99-

coefficients, and the associated value is the value of the wavelet coefficient. In the illustrated data example in table 182, the maximum value 184 is indicated to be 1000. Although a particularly advantageous compression technique for mass spectrometry data is shown, it will be appreciated that other compression techniques can be used. Although not preferred, the data may also be stored without compression.

In compressing the data according to a preferred embodiment, an intermediate format 186 is generated. The intermediate format 186 generally comprises a real number having a whole number portion 188 and a decimal portion 190. The whole number portion is the x-axis point 183 while the decimal portion is the value data 184 divided by the maximum data value. For example, in the data 182 a data value "25" is indicated at x-axis point "100". The intermediate value for this data point would be "100.025".

From the intermediate compressed data 186 the final compressed data 195 is generated. The first point of the intermediate data file becomes the starting point for the compressed data. Thereafter each data point in the compressed data 195 is calculated as follows: the whole number portion (left of the decimal) is replaced by the difference between the current and the last whole number. The remainder (right of the decimal) remains intact. For example, the starting point of the compressed data 195 is shown to be the same as the intermediate data point which is "100.025". The comparison between the first intermediate data point "100.025" and the second intermediate data point "150.220" is "50.220". Therefore, "50.220" becomes the second point of the compressed data 195. In a similar manner, the second intermediate point is "150.220" and the third intermediate data point is "500.0001". Therefore, the third compressed data becomes "350.000". The calculation for determining compressed data points is continued until the entire array of data points is converted to a single array of real numbers.

FIG. 42 generally describes the method of compressing mass spectrometry data, showing that the data file in block 201 is presented as an array of coefficients in block 202. The data starting point and maximum is determined as shown in block 203, and the intermediate real numbers are

-101-

As generally addressed above, the denoised, shifted, and baseline-corrected signal is not sufficiently processed for confidently calling the DNA composition of the biological sample. For example, although the baseline has generally been removed, there are still residual baseline effects present. These 5 residual baseline effects are therefore removed to increase the accuracy and confidence in making identifications.

To remove the residual baseline effects, FIG. 45 shows that the putative peaks 218, 222, and 224 are removed from the baseline corrected signal. The peaks are removed by identifying a center line 230, 232, and 234 of the 10 putative peaks 218, 222, and 224, respectively and removing an area to the left and to the right of the identified center line. For each putative peak, an area equal to twice the width (W) of the Gaussian is removed from the left of the center line, while an area equivalent to 50 daltons is removed from the right of the center line. It has been found that the area representing 50 daltons is 15 adequate to sufficiently remove the effect of salt adducts which may be associated with an actual peak. Such adducts appear to the right of an actual peak and are a natural effect from the chemistry involved in acquiring a mass spectrum. Although a 50 Dalton buffer has been selected, it will be appreciated that other ranges or methods can be used to reduce or eliminate adduct effects.

The peaks are removed and remaining minima 247 located as shown in FIG. 46 with the minima 247 connected to create signal 245. A quartic polynomial is applied to signal 245 to generate a residual baseline 250 as shown in FIG. 47. The residual baseline 250 is subtracted from the signal 225 to generate the final signal 255 as indicated in FIG. 48. Although the residual 25 baseline is the result of a quartic fit to signal 245, it will be appreciated that other techniques can be used to smooth or fit the residual baseline.

To determine peak height, as shown in FIG. 49, a Gaussian such as Gaussian 266, 268, and 270 is fit to each of the peaks, such as peaks 260, 262, and 264, respectively. Accordingly, the height of the Gaussian is 30 determined as height 272, 274, and 276. Once the height of each Gaussian peak is determined, then the method of identifying a biological compound 35 can move into the genotyping phase 65 as shown in FIG. 25.

-102-

An indication of the confidence that each putative peak is an actual peak can be discerned by calculating a signal-to-noise ratio for each putative peak. Accordingly, putative peaks with a strong signal-to-noise ratio are generally more likely to be an actual peak than a putative peak with a lower signal-to-noise

5 ratio. As described above and shown in FIG. 50, the height of each peak, such as height 272, 274, and 276, is determined for each peak, with the height being an indicator of signal strength for each peak. The noise profile, such as noise profile 97, is extrapolated into noise profile 280 across the identified peaks. At the center line of each of the peaks, a noise value is determined, such as noise

10 value 282, 283, and 284. With a signal values and a noise values generated, signal-to-noise ratios can be calculated for each peak. For example, the signal-to-noise ratio for the first peak in FIG. 50 would be calculated as signal value 272 divided by noise value 282, and in a similar manner the signal-to-noise ratio of the middle peak in FIG. 50 would be determined as signal 274 divided by

15 noise value 283.

Although the signal-to-noise ratio is generally a useful indicator of the presence of an actual peak, further processing has been found to increase the confidence by which a sample can be identified. For example, the signal-to-noise ratio for each peak in the preferred embodiment is preferably adjusted by

20 the goodness of fit between a Gaussian and each putative peak. It is a characteristic of a mass spectrometer that sample material is detected in a manner that generally complies with a normal distribution. Accordingly, greater confidence will be associated with a putative signal having a Gaussian shape than a signal that has a less normal distribution. The error resulting from having

25 a non-Gaussian shape can be referred to as a "residual error".

Referring to FIG. 51, a residual error is calculated by taking a root mean square calculation between the Gaussian 293 and the putative peak 290 in the data signal. The calculation is performed on data within one width on either side of a center line of the Gaussian. The residual error is calculated as:

30 where G is the Gaussian signal value, R is the putative peak value, and N is the number of points from -W to +W. The calculated residual error is used to generate an adjusted signal-to-noise ratio, as described below.

-103-

An adjusted signal noise ratio is calculated for each putative peak using the formula $(S/N) * EXP(-.1 * R)$, where S/N is the signal-to-noise ratio, and R is the residual error determined above. Although the preferred embodiment calculates an adjusted signal-to-noise ratio using a residual error for each peak, it
5 will be appreciated that other techniques can be used to account for the goodness of fit between the Gaussian and the actual signal.

Referring now to FIG. 52, a probability is determined that a putative peak is an actual peak. In making the determination of peak probability, a probability profile 300 is generated where the adjusted signal-to-noise ratio is the x-axis and
10 the probability is the y-axis. Probability is necessarily in the range between a 0% probability and a 100% probability, which is indicated as 1. Generally, the higher the adjusted signal-to-noise ratio, the greater the confidence that a putative peak is an actual peak.

At some target value for the adjusted signal-to-noise, it has been found
15 that the probability is 100% that the putative peak is an actual peak and can confidently be used to identify the DNA composition of a biological sample. However, the target value of adjusted signal-to-noise ratio where the probability is assumed to be 100% is a variable parameter which is to be set according to application specific criteria. For example, the target signal-to-noise ratio will be
20 adjusted depending upon trial experience, sample characteristics, and the acceptable error tolerance in the overall system. More specifically, for situations requiring a conservative approach where error cannot be tolerated, the target adjusted signal-to-noise ratio can be set to, for example, 10 and higher. Accordingly, 100% probability will not be assigned to a peak unless the adjusted
25 signal-to-noise ratio is 10 or over.

In other situations, a more aggressive approach may be taken as sample data is more pronounced or the risk of error may be reduced. In such a situation, the system may be set to assume a 100% probability with a 5 or greater target signal-to-noise ratio. Of course, an intermediate signal-to-noise
30 ratio target figure can be selected, such as 7, when a moderate risk of error can be assumed. Once the target adjusted signal-to-noise ratio is set for the method,

-104-

then for any adjusted signal-to-noise ratio a probability can be determined that a putative peak is an actual peak.

Due to the chemistry involved in performing an identification test, especially a mass spectrometry test of a sample prepared by DNA amplifications,

- 5 the allelic ratio between the signal strength of the highest peak and the signal strength of the second (or third and so on) highest peak should fall within an expected ratio. If the allelic ratio falls outside of normal guidelines, the preferred embodiment imposes an allelic ratio penalty to the probability. For example, FIG. 53 shows an allelic penalty 315 which has an x-axis 317 that is the ratio
- 10 between the signal strength of the second highest peak divided by signal strength of the highest peak. The y-axis 319 assigns a penalty between 0 and 1 depending on the determined allelic ratio. In the preferred embodiment, it is assumed that allelic ratios over 30% are within the expected range and therefore no penalty is applied. Between a ratio of 10% and 30%, the penalty is linearly
- 15 increased until at allelic ratios below 10% it is assumed the second-highest peak is not real. For allelic ratios between 10% and 30%, the allelic penalty chart 315 is used to determine a penalty 319, which is multiplied by the peak probability determined in FIG. 52 to determine a final peak probability. Although the preferred embodiment incorporates an allelic ratio penalty to account for a
- 20 possible chemistry error, it will be appreciated that other techniques may be used. Similar treatment will be applied to the other peaks.

With the peak probability of each peak determined, the statistical probability for various composition components may be determined. As an example, in order to determine the probability of each of three possible

- 25 combinations of two peaks, -- peak G, peak C and combinations GG, CC and GC. FIG. 54 shows an example where a most probable peak 325 is determined to have a final peak probability of 90%. Peak 325 is positioned such that it represents a G component in the biological sample. Accordingly, it can be maintained that there is a 90% probability that G exists in the biological sample.
- 30 Also in the example shown in FIG. 54, the second highest probability is peak 330 which has a peak probability of 20%. Peak 330 is at a position associated

-105-

with a C composition. Accordingly, it can be maintained that there is a 20% probability that C exists in the biological sample.

With the probability of G existing (90%) and the probability of C existing (20%) as a starting point, the probability of combinations of G and C existing
5 can be calculated. For example, FIG. 54 indicates that the probability of GG existing 329 is calculated as 72%. This is calculated as the probability of GG is equal to the probability of G existing (90%) multiplied by the probability of C not existing (100% -20%). So if the probability of G existing is 90% and the probability of C not existing is 80%, the probability of GG is 72%.

10 In a similar manner, the probability of CC existing is equivalent to the probability of C existing (20%) multiplied by the probability of G not existing (100% - 90%). As shown in FIG. 54, the probability of C existing is 20% while the probability of G not existing is 10%, so therefore the probability of CC is only 2%. Finally, the probability of GC existing is equal to the probability of G existing (90%) multiplied by the probability of C existing (20%). So if the
15 probability of G existing is 90% and the probability of C existing is 20%, the probability of GC existing is 18%. In summary form, then, the probability of the composition of the biological sample is:

probability of GG: 72%;
20 probability of GC: 18%; and
probability of CC: 2%.

Once the probabilities of each of the possible combinations has been determined, FIG. 55 is used to decide whether or not sufficient confidence exists to call the genotype. FIG. 55 shows a call chart 335 which has an x-axis 337
25 which is the ratio of the highest combination probability to the second highest combination probability. The y-axis 339 simply indicates whether the ratio is sufficiently high to justify calling the genotype. The value of the ratio may be indicated by M 340. The value of M is set depending upon trial data, sample composition, and the ability to accept error. For example, the value M may be
30 set relatively high, such as to a value 4 so that the highest probability must be at least four times greater than the second highest probability before confidence is established to call a genotype. However, if a certain level of error may be

-106-

acceptable, the value of M may be set to a more aggressive value, such as to 3, so that the ratio between the highest and second highest probabilities needs to be only a ratio of 3 or higher. Of course, moderate value may be selected for M when a moderate risk can be accepted. Using the example of FIG. 54, where
5 the probability of GG was 72% and the probability of GC was 18%, the ratio between 72% and 18% is 4.0, therefore, whether M is set to 3, 3.5, or 4, the system would call the genotype as GG. Although the preferred embodiment uses a ratio between the two highest peak probabilities to determine if a genotype confidently can be called, it will be appreciated that other methods
10 may be substituted. It will also be appreciated that the above techniques may be used for calculating probabilities and choosing genotypes (or more general DNA patterns) containing of combinations of more than two peaks.

Referring now to FIG. 56, a flow chart is shown generally defining the process of statistically calling genotype described above. In FIG. 56 block 402 shows that the height of each peak is determined and that in block 404 a noise profile is extrapolated for each peak. The signal is determined from the height of each peak in block 406 and the noise for each peak is determined using the noise profile in block 408. In block 410, the signal-to-noise ratio is calculated for each peak. To account for a non-Gaussian peak shape, a residual error is
20 determined in block 412 and an adjusted signal-to-noise ratio is calculated in block 414. Block 416 shows that a probability profile is developed, with the probability of each peak existing found in block 418. An allelic penalty may be applied in block 420, with the allelic penalty applied to the adjusted peak probability in block 422. The probability of each combination of components is
25 calculated in block 424 with the ratio between the two highest probabilities being determined in block 426. If the ratio of probabilities exceeds a threshold value then the genotype is called in block 428.

In another embodiment of the invention, the computing device 20 (Fig.
30 24) supports "standardless" genotyping by identifying data peaks that contain putative SNPs. Standardless genotyping is used, for example, where insufficient information is known about the samples to determine a distribution of expected

peak locations, against which an allelic penalty as described above can be reliably calculated. This permits the computing device to be used for identification of peaks that contain putative SNPs from data generated by any assay that fragments a targeted DNA molecule. For such standardless genotyping, peaks that are associated with an area under the data curve that deviates significantly from the typical area of other peaks in the data spectrum are identified and their corresponding mass (location along the x-axis) is determined.

More particularly, peaks that deviate significantly from the average area of other peaks in the data are identified, and the expected allelic ratio between data peaks is defined in terms of the ratio of the area under the data peaks. Theoretically, where each genetic loci has the same molar concentration of analyte, the area under each corresponding peak should be the same, thus producing a 1.0 ratio of the peak area between any two peaks. In accordance with the invention, peaks having a smaller ratio relative to the other peaks in the data will not be recognized as peaks. More particularly, peaks having an area ratio smaller than 30% relative to a nominal value for peak area will be assigned an allelic penalty. The mass of the remaining peaks (their location along the x-axis of the data) will be determined based on oligonucleotide standards.

Fig. 57 shows a flow diagram representation of the processing by the computing device 20 (Fig. 24) when performing standardless genotyping. In the first operation, represented by the flow diagram box numbered 502, the computing device receives data from the mass spectrometer. Next, the height of each putative peak in the data sample is determined, as indicated by the block 504. After the height of each peak in the mass spectrometer data is determined, a de-noise process 505 is performed, beginning with an extrapolation of the noise profile (block 506), followed by finding the noise of each peak (block 508) and calculating the signal to noise ratio for each data sample (block 510). Each of these operations may be performed in accordance with the description above for denoise operations 45 of Fig. 25. Other suitable denoise operations will occur to those skilled in the art.

-108-

The next operation is to find the residual error associated with each data point. This is represented by the block 512 in Figure 57. The next step, block 514, involves calculating an adjusted signal to noise ratio for each identified peak. A probability profile is developed next (block 516), followed by a 5 determination of the peak probabilities at block 518. In the preferred embodiment, the denoise operations of Fig. 57, comprising block 502 to block 518, comprise the corresponding operations described above in conjunction with Fig. 56 for block 402 through block 418, respectively.

The next action for the standardless genotype processing is to determine 10 an allelic penalty for each peak, indicated by the block 524. As noted above, the standardless genotype processing of Fig. 57 determines an allelic penalty by comparing area under the peaks. Therefore, rather than compare signal strength ratios to determine an allelic penalty, such as described above for Fig. 53, the standardless processing determines the area under each of the identified peaks 15 and compares the ratio of those areas. Determining the area under each peak may be computed using conventional numerical analysis techniques for calculating the area under a curve for experimental data.

Thus, the allelic penalty is assigned in accordance with Fig. 58, which shows that no penalty is assigned to peaks having a peak area relative to an 20 expected average area value that is greater than 0.30 (30%). The allelic penalty is applied to the peak probability value, which may be determined according to the process such as described in Fig. 52. It should be apparent from Fig. 58 that the allelic penalty imposed for peaks below a ratio of 30% is that such 25 peaks will be removed from further measurement and processing. Other penalty schemes, however, may be imposed in accordance with knowledge about the data being processed, as determined by those skilled in the art.

After the allelic penalty has been determined and applied, the standardless genotype processing compares the location of the remaining putative peaks to oligonucleotide standards to determine corresponding masses 30 in the processing for block 524. For standardless genotype data, the processing of the block 524 is performed to determine mass and genotype, rather than performing the operations corresponding to block 424, 426, and 428 of Fig. 33.

-109-

Techniques for performing such comparisons and determining mass will be known to those skilled in the art.

In another embodiment, the computing device 20 (Fig. 24) permits the detection and determination of the mass (location along the x-axis of the data) of

5 the sense and antisense strand of fragments generated in the assay. If desired, the computing device may also detect and determine the quantity (area under each peak) of the respective sense and antisense strands, using a similar technique to that described above for standardless genotype processing. The data generated for each type of strand may then be combined to achieve a data

10 redundancy and to thereby increase the confidence level of the determined genotype. This technique obviates primer peaks that are often observed in data from other diagnostic methods, thereby permitting a higher level of multiplexing. In addition, when quantitation is used in pooling experiments, the ratio of the measured peak areas is more reliably calculated than the peak identifying

15 technique, due to data redundancy.

Fig. 23 is a flow diagram that illustrates the processing implemented by the computing device 20 to perform sense and antisense processing. In the first operation, represented by the flow diagram box numbered 602, the computing device receives data from the mass spectrometer. This data will include data for

20 the sense strand and antisense strand of assay fragments. Next, the height of each putative peak in the data sample is determined, as indicated by the block 604. After the height of each peak in the mass spectrometer data is determined, a de-noise process 605 is performed, beginning with an operation that extrapolates the noise profile (block 606), followed by finding the noise of

25 each peak (block 608) and calculating the signal to noise ratio for each data sample (block 610). Each of these operations may be performed in accordance with the description above for the denoise operations 45 of Fig. 25. Other suitable denoise operations will occur to those skilled in the art. The next operation is to find the residual error associated with each data point. This is

30 represented by the block 612 in Figure 36.

After the residual error for the data of the sense strand and antisense strand has been performed, processing to identify the genotypes will be

-110-

performed for the sense strand and also for the antisense strand. Therefore, Fig. 23 shows that processing includes sense strand processing (block 630) and antisense strand processing (block 640). Each block 630, 640 includes processing that corresponds to adjusting the signal to noise ratio, developing a 5 probability profile, determining an allelic penalty, adjusting the peak probability by the allelic penalty, calculating genotype probabilities, and testing genotype probability ratios, such as described above in conjunction with blocks 414 through 426 of Fig. 56. The processing of each block 630, 640 may, if desired, include standardless processing operations such as described above in 10 conjunction with Fig. 57. The standardless processing may be included in place of or in addition to the processing operations of Fig. 56.

After the genotype probability processing is completed, the data from the sense strand and antisense strand processing is combined and compared to expected database values to obtain the benefits of data redundancy as between 15 the sense strand and antisense strand. Those skilled in the art will understand techniques to take advantage of known data redundancies between a sense strand and antisense strand of assay fragments. This processing is represented by the block 650. After the data from the two strands is combined for processing, the genotype processing is performed (block 660) and the genotype 20 is identified.

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

-111-

WHAT IS CLAIMED IS:

1. A subcollection of samples from a target population, comprising:
 - a plurality of samples, wherein the samples are selected from the group consisting of blood, tissue, body fluid, cell, seed, microbe, pathogen and reproductive tissue samples; and
 - a symbology on the containers containing the samples, wherein the symbology is representative of the source and/or history of each sample, wherein:
 - the target population is a healthy population that has not been selected
- 5 for any disease state;
 - the collection comprises samples from the healthy population; and
 - the subcollection is obtained by sorting the collection according to specified parameters.
- 10 2. The subcollection of claim 1, wherein the parameters are selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence
- 15 20 of a disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and exposure to environmental factors.
- 25 3. The subcollection of claim 1, wherein the symbology is a bar code.
4. A method of producing a database, comprising:
 - 25 identifying healthy members of a population;
 - obtaining data comprising identifying information and obtaining historical information and data relating to the identified members of the population and their immediate family;
 - entering the data into a database for each member of the population and
- 30 30 associating the member and the data with an indexer.
5. The method of claim 4, further comprising:
 - obtaining a body tissue or body fluid sample;

-112-

analyzing the body tissue or body fluid in the sample; and
entering the results of the analysis for each member into the database
and associating each result with the indexer representative of each member.

6. A database produced by the method of claim 4.
- 5 7. A database produced by the method of claim 5.
8. A database, comprising:
 - datapoints representative of a plurality of healthy organisms from whom biological samples are obtained,
wherein each datapoint is associated with data representative of
 - 10 the organism type and other identifying information.
 9. The database of claim 8, wherein the datapoints are answers to questions regarding one or more of a parameters selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of a disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and exposure to environmental factors.
 10. The database of claim 9, wherein the organisms are mammals and the samples are body fluids or tissues.
 11. The database of claim 9, wherein the samples are selected from blood, blood fractions, cells and subcellular organelles.
 - 25 12. The database of claim 8, further comprising, phenotypic data from an organism.
 13. The database of claim 12, wherein the data includes one of physical characteristics, background data, medical data, and historical data.
 14. The database of claim 8, further comprising,
 - 30 genotypic data from nucleic acid obtained from an organism.

-113-

15. The database of claim 14, wherein genotypic data includes, genetic markers, non-coding regions, microsatellites, RFLPs, VNTRs, historical data of the organism, medical history, and phenotypic information.

16. The database of claim 8 that is a relational database.

5 17. The database of claim 16, wherein the data are related to an indexer datapoint representative of each organism from whom data is obtained.

18. A method of identifying polymorphisms that are candidate genetic markers, comprising:

identifying a polymorphism; and

10 identifying any pathway or gene linked to the locus of the polymorphism, wherein

the polymorphisms are identified in samples associated with a target population that comprises healthy subjects.

19. The method of claim 18, wherein the polymorphism is identified by 15 detecting the presence of target nucleic acids in a sample by a method, comprising the steps of:

a) hybridizing a first oligonucleotide to the target nucleic acid;

b) hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;

20 c) ligating the hybridized oligonucleotides; and

c) detecting hybridized first oligonucleotide by mass spectrometry as an indication of the presence of the target nucleic acid.

20. The method of claim 18, wherein the polymorphism is identified by detecting target nucleic acids in a sample by a method, comprising the steps of:

25 a) hybridizing a first oligonucleotide to the target nucleic acid and hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;

b) contacting the hybridized first and second oligonucleotides with a cleavage enzyme to form a cleavage product; and

30 c) detecting the cleavage product by mass spectrometry as an indication of the presence of the target nucleic acid.

-114-

21. The method of claim 20 wherein the samples are from subjects in a healthy database.

22. The method of claim 18, wherein the polymorphism is identified by identifying target nucleic acids in a sample by primer oligo base extension
5 (probe).

23. The method of 22, wherein primer oligo base extension, comprises:

- a) obtaining a nucleic acid molecule that contains a target nucleotide;
- b) optionally immobilizing the nucleic acid molecule onto a solid support,
10 to produce an immobilized nucleic acid molecule;
- c) hybridizing the nucleic acid molecule with a primer oligonucleotide that is complementary to the nucleic acid molecule at a site adjacent to the target nucleotide;
- d) contacting the product of step c) with a composition comprising a
15 dideoxynucleoside triphosphate or a 3'-deoxynucleoside triphosphates and a polymerase, so that only a dideoxynucleoside or 3'-deoxynucleoside triphosphate that is complementary to the target nucleotide is extended onto the primer; and
- e) detecting the extended primer, thereby identifying the target nucleotide.

20 24. The method of claim 23, wherein detection of the extended primer is effected by mass spectrometry, comprising:

- ionizing and volatilizing the product of step d); and
- detecting the extended primer by mass spectrometry, thereby identifying the target nucleotide.

25 25. The method of claim 24, wherein;
samples are presented to the mass spectrometer as arrays on chips; and
each sample occupies a volume that is about the size of the laser spot projected by the laser in a mass spectrometer used in matrix-assisted laser desorption/ionization (MALDI) spectrometry.

-115-

26. A combination, comprising:

a database containing parameters associated with a datapoint

representative of a subject from whom samples are obtained, wherein the subjects are healthy; and

5 an indexed collection of the samples, wherein the index identifies the subject from whom the sample was obtained.

27 The combination of claim 26, wherein the parameter is selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of
10 physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and
15 exposure to environmental factors.

28. The combination of claim 26, wherein the database further contains genotypic data for each subject.

29. The combination of claim 26, wherein the samples are blood.

30 A data storage medium, comprising the database of claim 8.

20 31. A computer system, comprising the database of claim 8.

32. A system for high throughput processing of biological samples, comprising:

a process line comprising a plurality of processing stations, each of which performs a procedure on a biological sample contained in a
25 reaction vessel;

a robotic system that transports the reaction vessel from processing station to processing station;

a data analysis system that receives test results of the process line and automatically processes the test results to make a determination
30 regarding the biological sample in the reaction vessel;

a control system that determines when the test at each processing station is complete and, in response, moves the reaction vessel to

-116-

the next test station, and continuously processes reaction vessels one after another until the control system receives a stop instruction; and

5 a database of claim 8, wherein the samples tested by the automated process line comprise samples from subjects in the database.

33. The system of claim 32, wherein one of the processing stations comprises a mass spectrometer.

34. The system of claim 32, wherein the data analysis system processes the test results by receiving test data from the mass spectrometer 10 such that the test data for a biological sample contains one or more signals, whereupon the data analysis system determines the area under the curve of each signal and normalizes the results thereof and obtains a substantially quantitative result representative of the relative amounts of components in the tested sample.

15 35. A method for high throughput processing of biological samples, the method comprising:

transporting a reaction vessel along a system of claim 32, comprising a process line having a plurality of processing stations, each of which performs a procedure on one or more biological samples contained in the reaction vessel;
20 determining when the test procedure at each processing station is complete and, in response, moving the reaction vessel to the next processing station;
receiving test results of the process line and automatically processing the 25 test results to make a data analysis determination regarding the biological samples in the reaction vessel; and
processing reaction vessels continuously one after another until receiving a stop instruction, wherein the samples tested by the automated process line comprise samples from subjects in the database.

30 36. The method of 35, wherein one of the processing stations comprises a mass spectrometer.

-117-

37. The method of claim 36, wherein the samples are analyzed by a method comprising primer oligo base extension (probe).

38. The method of claim 37; further comprising:
processing the test results by receiving test data from the mass
5 spectrometer such that the test data for a biological sample contains one or
more signals or numerical values representative of signals, whereupon the data
analysis system determines the area under the curve of each signal and
normalizes the results thereof and obtains a substantially quantitative result
representative of the relative amounts of components in the tested sample.

10 39. The method of claim 37, wherein primer oligo base extension,
comprises:
a) obtaining a nucleic acid molecule that contains a target nucleotide;
b) optionally immobilizing the nucleic acid molecule onto a solid support,
to produce an immobilized nucleic acid molecule;
15 c) hybridizing the nucleic acid molecule with a primer oligonucleotide that
is complementary to the nucleic acid molecule at a site adjacent to the target
nucleotide;
d) contacting the product of step c) with composition comprising a
dideoxynucleoside triphosphate or a 3'-deoxynucleoside triphosphates and a
20 polymerase, so that only a dideoxynucleoside or 3'-deoxynucleoside triphosphate
that is complementary to the target nucleotide is extended onto the primer; and
e) detecting the primer, thereby identifying the target nucleotide.

40. The method of 39, wherein detection of the extended primer is
effected by mass spectrometry, comprising:
25 ionizing and volatizing the product of step d); and
detecting the extended primer by mass spectrometry, thereby identifying
the target nucleotide.

41. The method of claim 36, wherein the target nucleic acids in the
sample are detected and/or identified by a method, comprising the steps of:
30 a) hybridizing a first oligonucleotide to the target nucleic acid;
b) hybridizing a second oligonucleotide to an adjacent region of the
target nucleic acid;

-118-

- c) ligating then hybridized oligonucleotides; and
- c) detecting hybridized first oligonucleotide by mass spectrometry as an indication of the presence of the target nucleic acid.

42. The method of claim 36, wherein the target nucleic acids in the sample are detected and/or identified by a method, comprising the steps of:

- a) hybridizing a first oligonucleotide to the target nucleic acid and hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;
- b) contacting the hybridized first and second oligonucleotides with a cleavage enzyme to form a cleavage product; and
- c) detecting the cleavage product by mass spectrometry as an indication of the presence of the target nucleic acid.

43. A method of producing a database stored in a computer memory, comprising:

- 15 identifying healthy members of a population;
- obtaining identifying and historical information and data relating to the identified members of the population;
- entering the member-related data into the computer memory database for each identified member of the population and associating the member and the data with an indexer.

44. The method of claim 43, further comprising:

- obtaining a body tissue or body fluid sample of an identified member;
 - analyzing the body tissue or body fluid in the sample; and
 - entering the results of the analysis for each member into the computer
- 25 memory database and associating each result with the indexer representative of each member.

45. A database produced by the method of claim 43.

46. A database produced by the method of claim 44.

47. The database of claim 8, wherein:

- 30 the organisms are selected from among animals, bacteria, fungi, protozoans and parasites and

-119-

each datapoint is associated with parameters representative of the organism type and identifying information.

48. The database of claim 43, further comprising, phenotypic data regarding each subject.

5 49. The database of claim 47 that is a relational database and the parameters are the answers to the questions in the questionnaire.

50. The database of claim 8, further comprising,

genotypic data of nucleic acid of the subject, wherein genotypic data includes, but is not limited to, genetic markers, non-coding regions,

10 microsatellites, restriction fragment length polymorphisms (RFLPs), variable number tandem repeats (VNTRs), historical day of the organism, the medical history of the subject, phenotypic information, and other information.

51. A database, comprising data records stored in computer memory, wherein the data records contain information that identifies healthy members of 15 a population, and also contain identifying and historical information and data relating to the identified members.

52. The database of claim 51, further comprising an index value for each identified member that associates each member of the population with the identifying and historical information and data.

20 53. A computer system, comprising the database of claim 51.

54. An automated process line, comprising the database of claim 51.

55. A method for determining a polymorphism that correlates with age, ethnicity or gender, comprising:

identifying a polymorphism; and

25 determining the frequency of the polymorphism with increasing age, with ethnicity or with gender in a healthy population.

56. A method for determining whether a polymorphism correlates with susceptibility to morbidity, early mortality, or morbidity and early mortality, comprising;

30 identifying a polymorphism; and

determining the frequency of the polymorphism with increasing age in a healthy population.

-120-

57. A high throughput method of determining frequencies of genetic variations, comprising:

selecting a healthy target population and a genetic variation to be assessed;

5 pooling a plurality of samples of biopolymers obtained from members of the population,

determining or detecting the biopolymer that comprises the variation by mass spectrometry;

obtaining a mass spectrum or a digital representation thereof; and

10 determining the frequency of the variation in the population.

58. The method of claim 57, wherein:

the variation is selected from the group consisting of an allelic variation, a post-translational modification, a nucleic modification, a label, a mass modification of a nucleic acid and methylation; and/or

15 the biopolymer is a nucleic acid, a protein, a polysaccharide, a lipid, a small organic metabolite or intermediate, wherein the concentration of biopolymer of interest is the same in each of the samples; and/or

the frequency is determined by assessing the method comprising determining the area under the peak in the mass spectrum or digital

20 representation thereof corresponding to the mass of the biopolymer comprising the genomic variation.

59. The method of claim 58, wherein the method for determining the frequency is effected by determining the ratio of the signal or the digital representation thereof to the total area of the entire mass spectrum, which is

25 corrected for background.

60. A method for discovery of a polymorphism in a population, comprising:

sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;

30 isolating a biopolymer from each identified sample;

optionally pooling each isolated biopolymer;

optionally amplifying the amount of biopolymer;

-121-

cleaving the pooled biopolymers to produce fragments thereof;
obtaining a mass spectrum of the resulting fragments and comparing the mass spectrum with a control mass spectrum to identify differences between the spectra and thereby identifying any polymorphisms; wherein:

5 the control mass spectrum is obtained from unsorted samples in the collection or samples sorted according to a different parameter.

61. The method of claim 60, wherein cleaving is effected by contacting the biopolymer with an enzyme.

62. The method of claim 61, wherein the enzyme is selected from the 10 group consisting of nucleotide glycosylase, a nickase and a type IIS restriction enzyme.

63. The method of claim 60, wherein the biopolymer is a nucleic acid or a protein.

64. The method of claim 60, wherein the mass spectrometric 15 format is selected from among Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof.

65. A method for discovery of a polymorphism in a population, comprising:

20 obtaining samples of body tissue or fluid from a plurality of organisms; isolating a biopolymer from each sample; pooling each isolated biopolymer;

optionally amplifying the amount of biopolymer;

cleaving the pooled biopolymers to produce fragments thereof;

25 obtaining a mass spectrum of the resulting fragments; comparing the frequency of each fragment to identify fragments present in amounts lower than the average frequency, thereby identifying any polymorphisms.

66. The method of claim 65, wherein cleaving is effected by contacting 30 the biopolymer with an enzyme.

-122-

67. The method of claim 66, wherein the enzyme is selected from the group consisting of nucleotide glycosylase, a nickase and a type IIS restriction enzyme.

68. The method of claim 65, wherein the biopolymer is a nucleic acid
5 or a protein.

69. The method of claim 65, wherein the mass spectrometric format is selected from among Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof.

10 70. The method of claim 65, wherein the samples are obtained from healthy subjects.

71. A method of correlating a polymorphism with a parameter, comprising:

sorting the database of claim 8 according to a selected parameter to
15 identify samples that match the selected parameter;
isolating a biopolymer from each identified sample;
pooling each isolated biopolymer;
optionally amplifying the amount of biopolymer;
determining the frequency of the polymorphism in the pooled
20 biopolymers, wherein:
an alteration of the frequency of the polymorphism compared to a control,
indicates a correlation of the polymorphism with the selected parameter; and
the control is the frequency of the polymorphism in pooled biopolymers
obtained from samples identified from an unsorted database or from a database
25 sorting according to a different parameter.

72. The method of claim 71, wherein the parameter is selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence
30 of a disease in immediate family (parent, siblings, children), use of prescription

-124-

81. A method for haplotyping polymorphisms in a nucleic acid, comprising:

- (a) sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;
- 5 (b) isolating nucleic acid from each identified sample;
- (c) optionally pooling each isolated nucleic acid;
- (d) amplifying the amount of nucleic acid;
- (e) forming single-stranded nucleic acid and splitting each single-strand into a separate reaction vessel;
- 10 (f) contacting each single-stranded nucleic acid with an adaptor nucleic acid to form an adaptor complex;
- (g) contacting the adaptor complex with a nuclease and a ligase;
- (h) contacting the products of step (g) with a mixture that is capable of amplifying a ligated adaptor to produce an extended product;
- 15 (i) obtaining a mass spectrum of each nucleic acid resulting from step (h) and detecting a polymorphism by identifying a signal corresponding to the extended product;
- (j) repeating steps (f) through (i) utilizing an adaptor nucleic acid able to hybridize with another adapter nucleic acid that hybridizes to a different sequence on the same strand; whereby
- 20 the polymorphisms are haplotyped by detecting more than one extended product.

82. The method of claim 1, wherein the nuclease is FEN-1.

83. A method for haplotyping polymorphisms in a population, comprising:

- sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;
- isolating a nucleic acid from each identified sample;
- pooling each isolated nucleic acid;
- 30 optionally amplifying the amount of nucleic acid;
- contacting the nucleic acid with at least one enzyme to produce fragments thereof;

-126-

- using the putative peaks to generate a residual baseline;
removing the residual baseline from the intermediate data set to generate
a corrected data set;
locating, responsive to removing the residual baseline, a probable peak in
5 the
corrected data set; and
identifying, using the located probable peak, the biological sample;
wherein the generated biological sample data set comprises data from
sense
10 strands and antisense strands of assay fragments.
89. The method according to claim 88, wherein identifying includes
combining
data from the sense strands and the antisense strands, and comparing the data
against expected sense strand and antisense strand values, to identify the
15 biological
sample.
90. The method according to claim 88, wherein identifying includes
deriving a peak probability for the probable peak, in accordance with whether the
probable peak is from sense strand data or from antisense strand data.
- 20 91. The method according to claim 88, wherein identifying includes
deriving a peak probability for the probable peak and applying an allelic penalty in
response to a
ratio between a calculated area under the probable peak and a calculated
expected average area under all peaks in the data set.
- 25 92. A method for identifying a biological sample, comprising:
generating a data set indicative of the composition of the biological
sample;
denoising the data set to generate denoised data;
deleting the baseline from the denoised data to generate an intermediate
30 data
set;

-127-

- defining putative peaks for the biological sample; using the putative peaks to generate a residual baseline; removing the residual baseline from the intermediate data set to generate
- a
- 5 corrected data set;
- locating, responsive to removing the residual baseline, a probable peak in the corrected data set; and
- identifying, using the located probable peak, the biological sample; wherein identifying includes deriving a peak probability for the probable
- 10 peak and
- applying an allelic penalty in response to a ratio between a calculated area under the probable peak and a calculated expected average area under all peaks in the data set.
- 15 93. The method according to claim 92, wherein identifying includes comparing
- data from probable peaks that did not receive an applied allelic penalty to determine their mass in accordance with oligonucleotide biological data.
94. The method according to claim 92, wherein the allelic penalty is
- 20 not applied to probable peaks whose ratio of area under the peak to the expected area value is greater than 30%.
95. A method for detecting a polymorphism in a nucleic acid, comprising:
- amplifying a region of the nucleic acid to produce an amplicon, wherein
- 25 the resulting amplicon comprises one or more enzyme restriction sites;
- contacting the amplicon with a restriction enzyme to produce fragments;
- obtaining a mass spectrum of the resulting fragments and analyzing signals in the mass spectrum by the method of claim 88; whereby:
- the polymorphism is detected from the pattern of the signals.
- 30 96. A subcollection of samples from a target population, comprising:
- a plurality of samples, wherein the samples are selected from the group consisting of nucleic acids, fetal tissue, protein samples; and

-128-

a symbology on the containers containing the samples, wherein the symbology is representative of the source and/or history of each sample, wherein:

the target population is a healthy population that has not been selected
5 for any disease state;

the collection comprises samples from the healthy population; and
the subcollection is obtained by sorting the collection according to
specified parameters.

97. The combination of claim 26, wherein the samples are selected
10 selected from the group consisting of nucleic acids, fetal tissue, protein, tissue,
body fluid, cell, seed, microbe, pathogen and reproductive tissue samples.

98. A combination, comprising the database of claim 8 and a mass
spectrometer.

99. The combination of claim 98 that is an automated process line for
15 analyzing biological samples.

100. A system for high throughput processing of biological samples,
comprising:

20 a database of claim 8, wherein the samples tested by the automated
process line comprise samples from subjects in the database; and
a mass spectrometry for analysis of biopolymers in the samples.

DNA Bank

Number of Samples	3912
-------------------	------

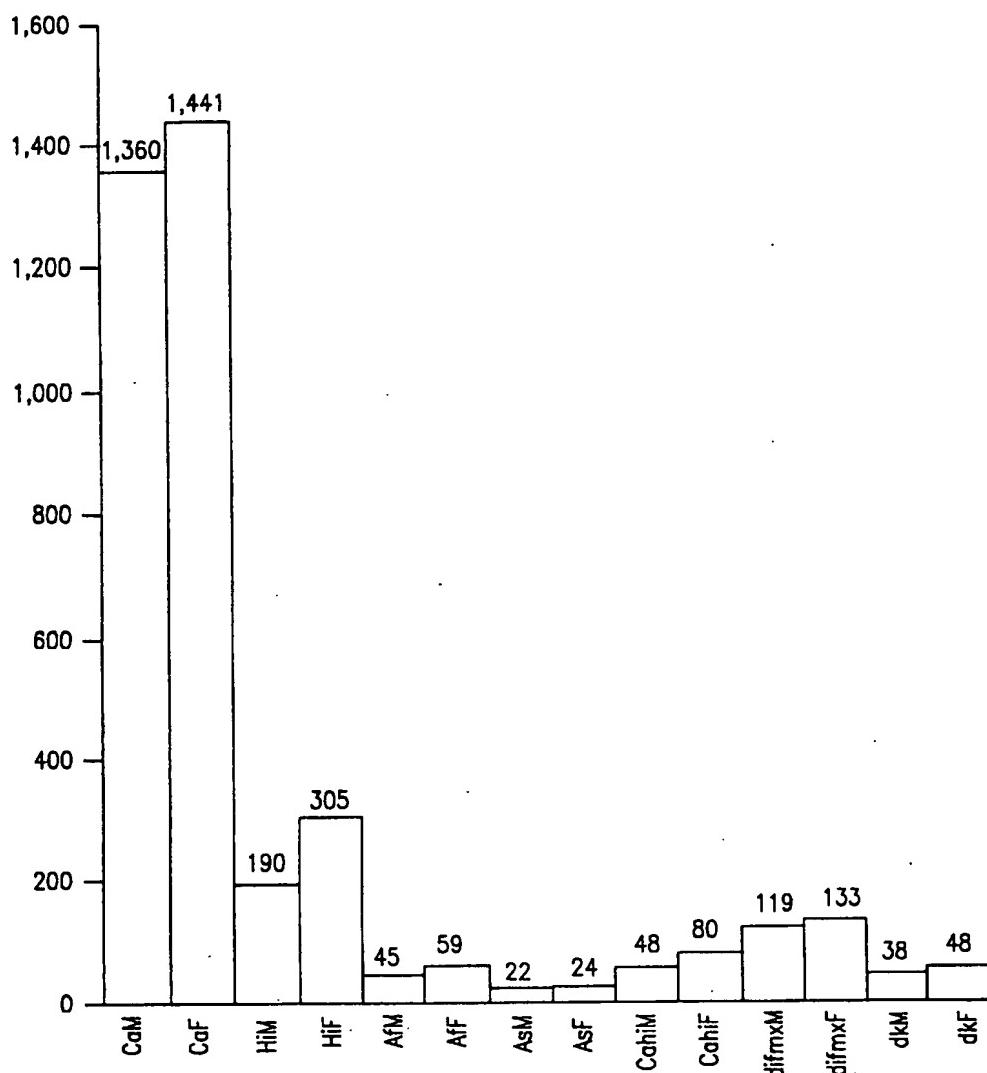


FIG. IA

2 / 51

Caucasians

Number of Samples	2801
-------------------	------

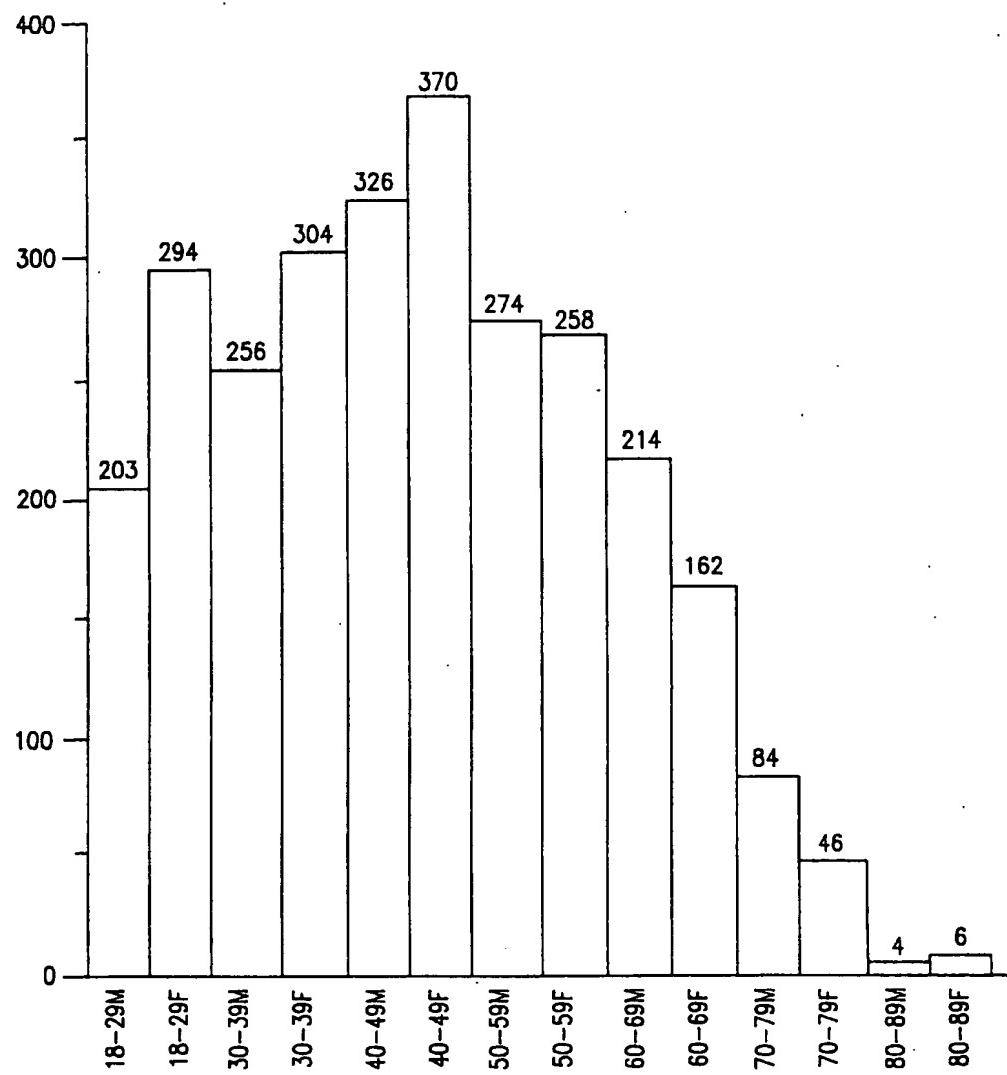


FIG. IB

3 / 51

Hispanics

Number of Samples	495
-------------------	-----

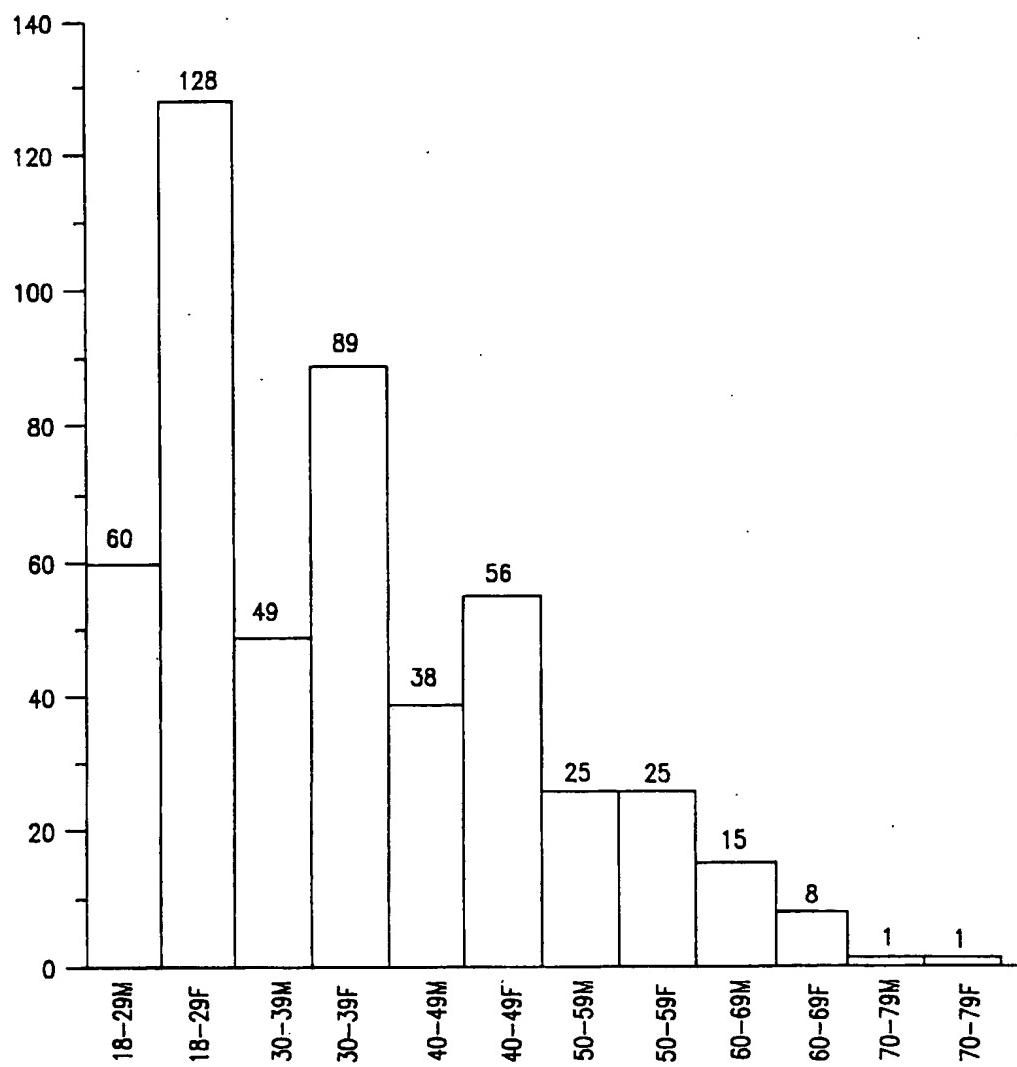
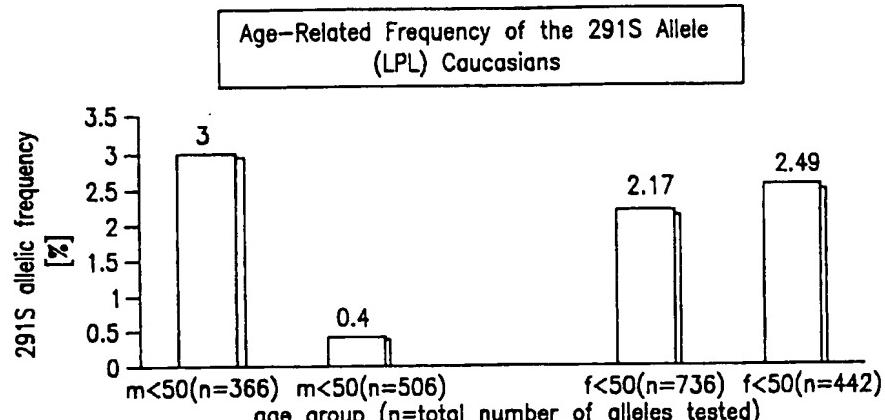
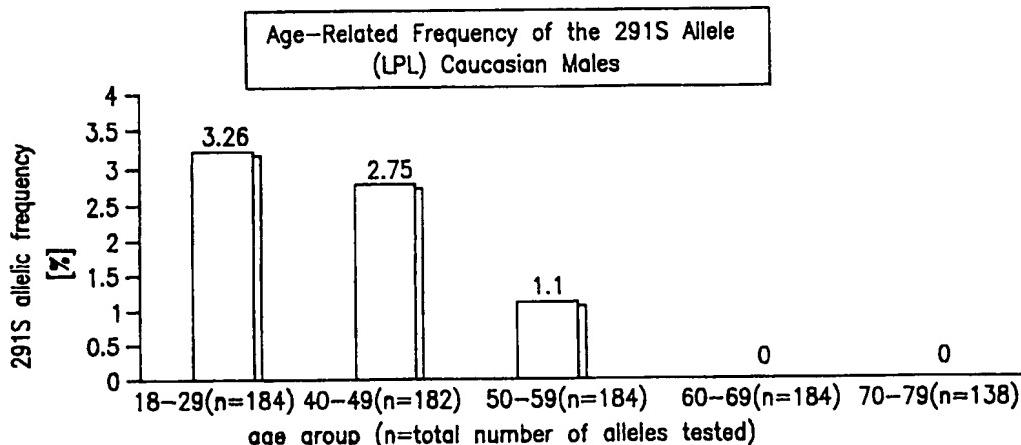


FIG. IC

4 / 51



age- and sex-distribution of the 291S allele of the lipoprotein lipase gene. A total of 436 males and 586 females were investigated.

FIG. 2A

Age-related distribution of the 291S allele of the lipoprotein lipase gene within the male Caucasian population. A total of 436 males were tested.

FIG. 2B

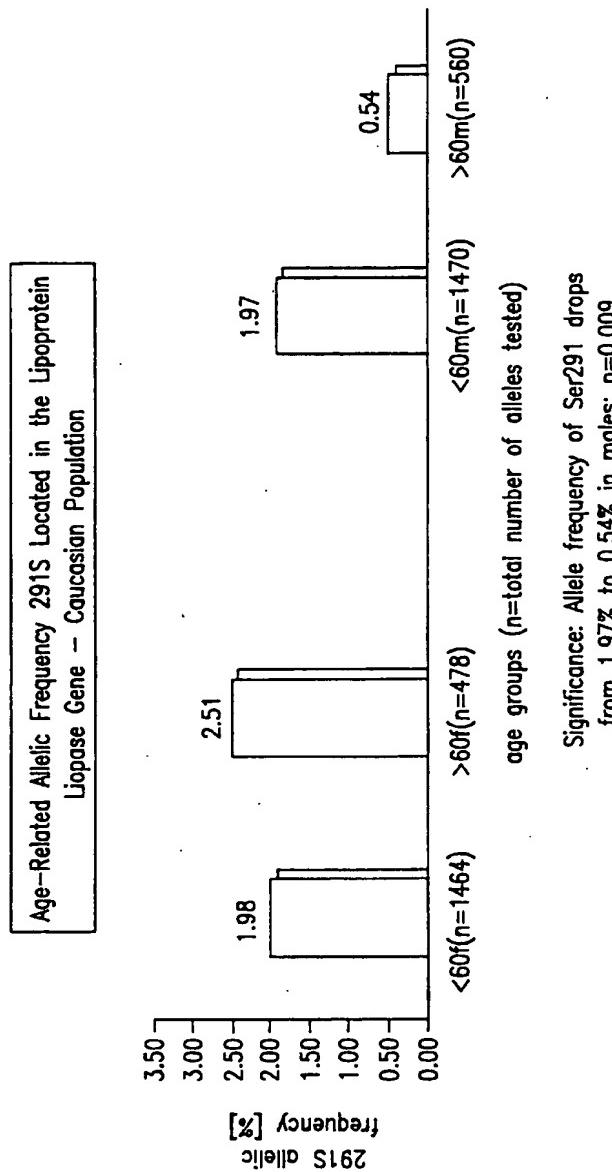


FIG. 2C

6/51

Questionnaire for
Population-Based
Sample Banking

Data Collection Form

Collection Information

Consent Form Signed Yes No

Date of Collection (MM/DD/YY) ____/____/98

Time of Sample Collection(nearest hour in 24 hour clock format) _____

Initials of Data Collector _____ Collecting Agency _____

Affix Barcode Here

(DO NOT COMPLETE: (For Date Entry Only) Sample _____ intact _____ lost _____ broken

Donor information

Sex: Male Female Date of Birth (MM/YY) ____/____/____

In which state do you live? _____ How long have you lived there? _____ Years

What is your highest grade you completed in school?

- less than 8th grade 8th,9th,10th or 11th grade high school graduate or equivalency
 some college 2 yr degree college graduate 4 yr degree post graduate education or degree

To the best of your knowledge what is the Ethnic Origin of your:

Father Mother

- | | |
|--------------------------|---|
| <input type="checkbox"/> | Caucasian (please check specific geographic area below if known) |
| <input type="checkbox"/> | Northern Europe (Austria,Denmark,Finland,France,Germany,Netherlands,Norway,Sweden,Switzerland,U.K.) |
| <input type="checkbox"/> | Southern Europe (Greece,Italy,Spain) |
| <input type="checkbox"/> | Eastern Europe (Czechoslovakia,Hungary,Poland,Russia,Yugoslavia) |
| <input type="checkbox"/> | Middle Eastern (Israel,Egypt,Iran,Iraq,Jordan,Syria, other Arab States) |
| <input type="checkbox"/> | African-American |
| <input type="checkbox"/> | Hispanic (please check specific geographic area below if known) |
| <input type="checkbox"/> | Mexico |
| <input type="checkbox"/> | Central America,South American |
| <input type="checkbox"/> | Cuba,Puerto Rico, other Caribbean |
| <input type="checkbox"/> | Asian (please check specific geographic area below if known) |
| <input type="checkbox"/> | Japanese |
| <input type="checkbox"/> | Chinese |
| <input type="checkbox"/> | Korean |
| <input type="checkbox"/> | Vietnamese |
| <input type="checkbox"/> | other Asian |
| <input type="checkbox"/> | Other _____ |
| <input type="checkbox"/> | Don't know |

Health information: Have you or has anyone in your immediate family(parents,brothers,sisters, or your children) had the following? Check all that apply

Disease: You Mother Father Sister Brother Child

Heart Disease Stroke or Arteriosclerosis
Cancer (Specify type if known)
Alzheimer's Disease or Dementia
Chronic inflammatory or Autoimmune Disease
Nervous System Disease like Multiple Sclerosis
Other (please specify)

Additional health information details you would like to provide:

FIG. 3

SUBSTITUTE SHEET (RULE 26)

7/51

Sample Banks

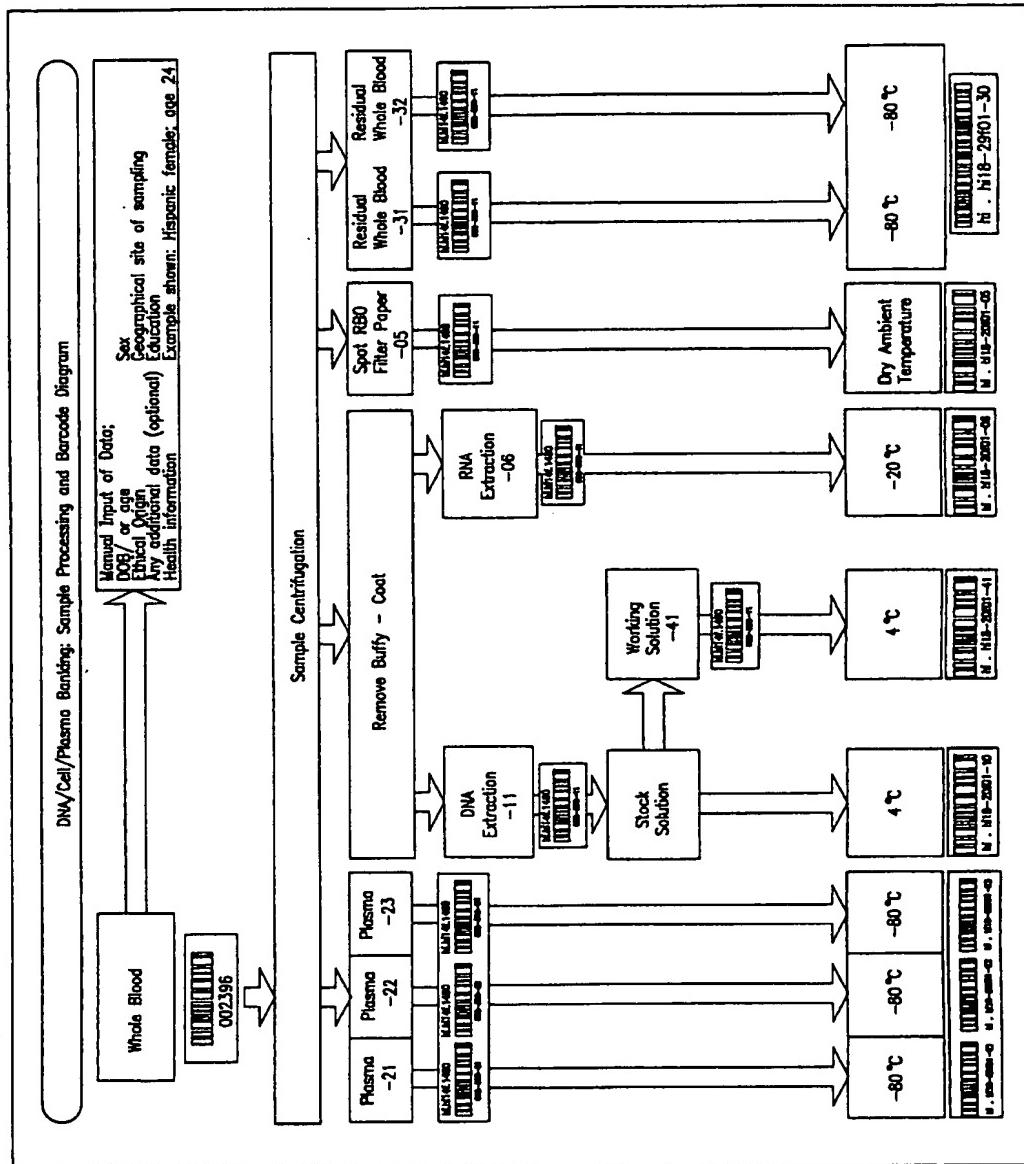


FIG. 4

8/51

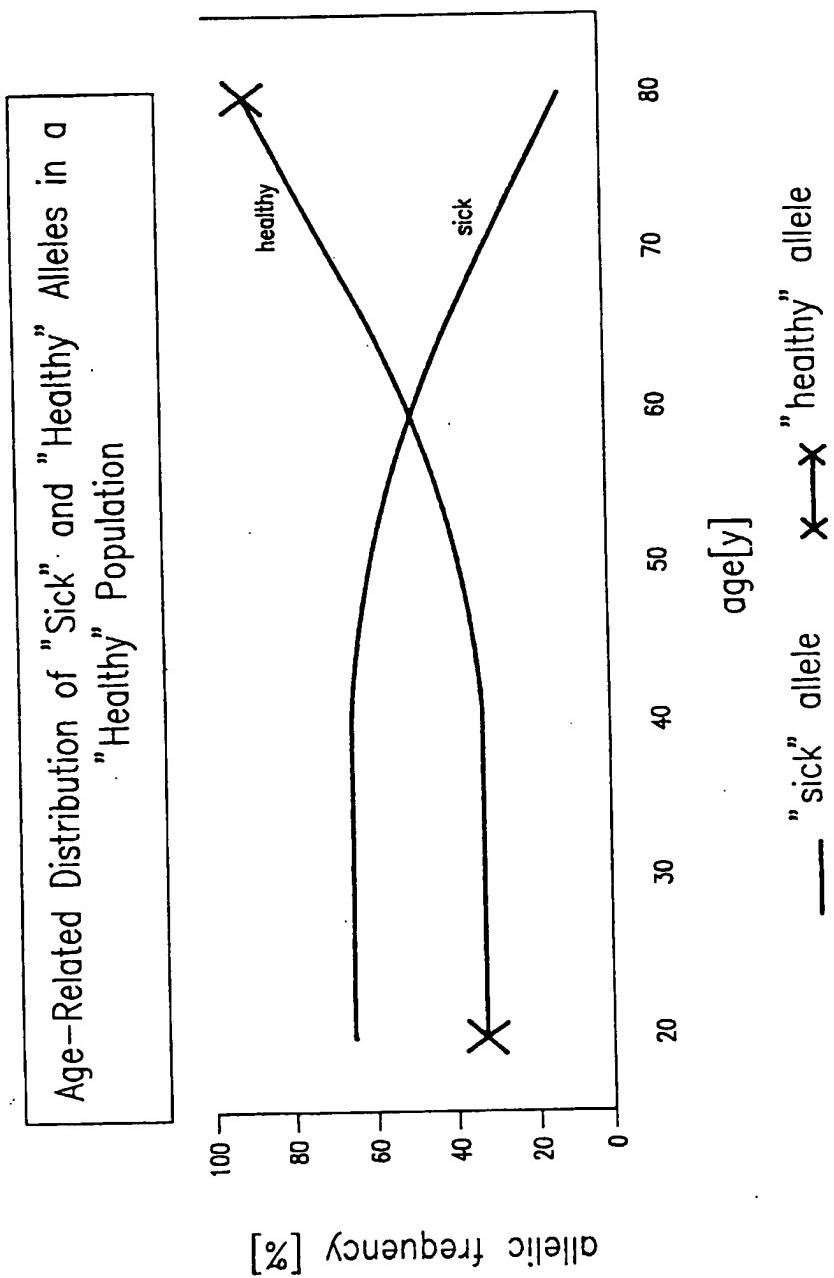


FIG. 5

9 / 51

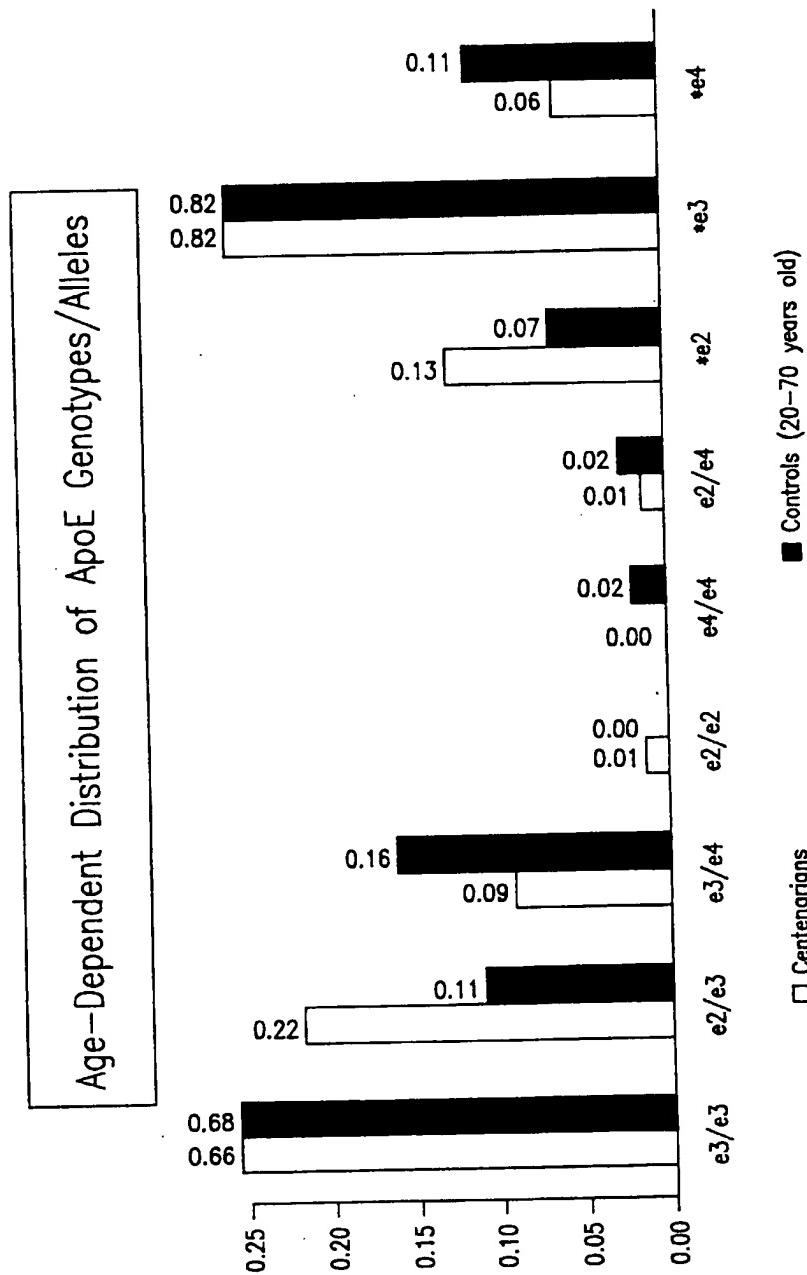
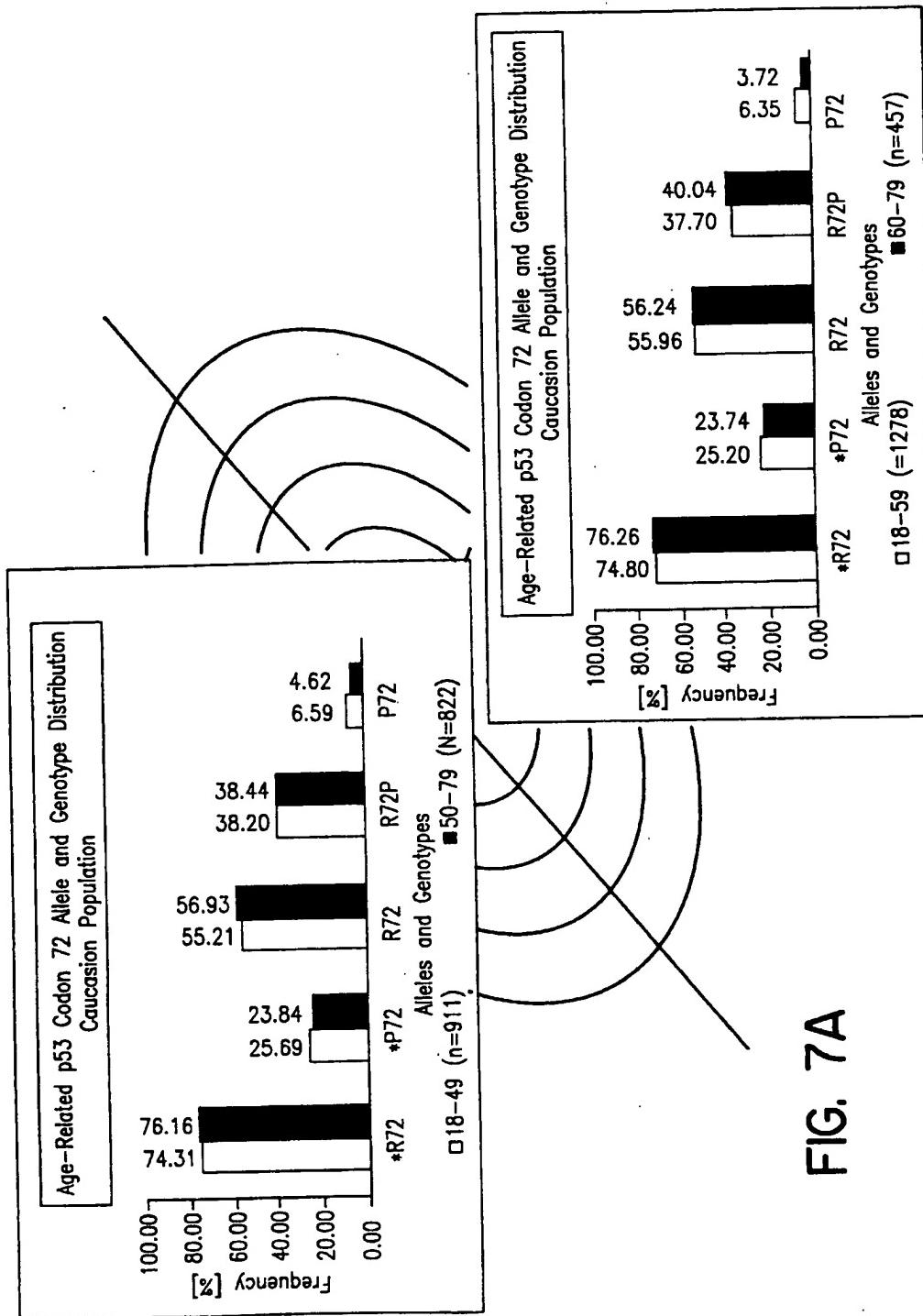
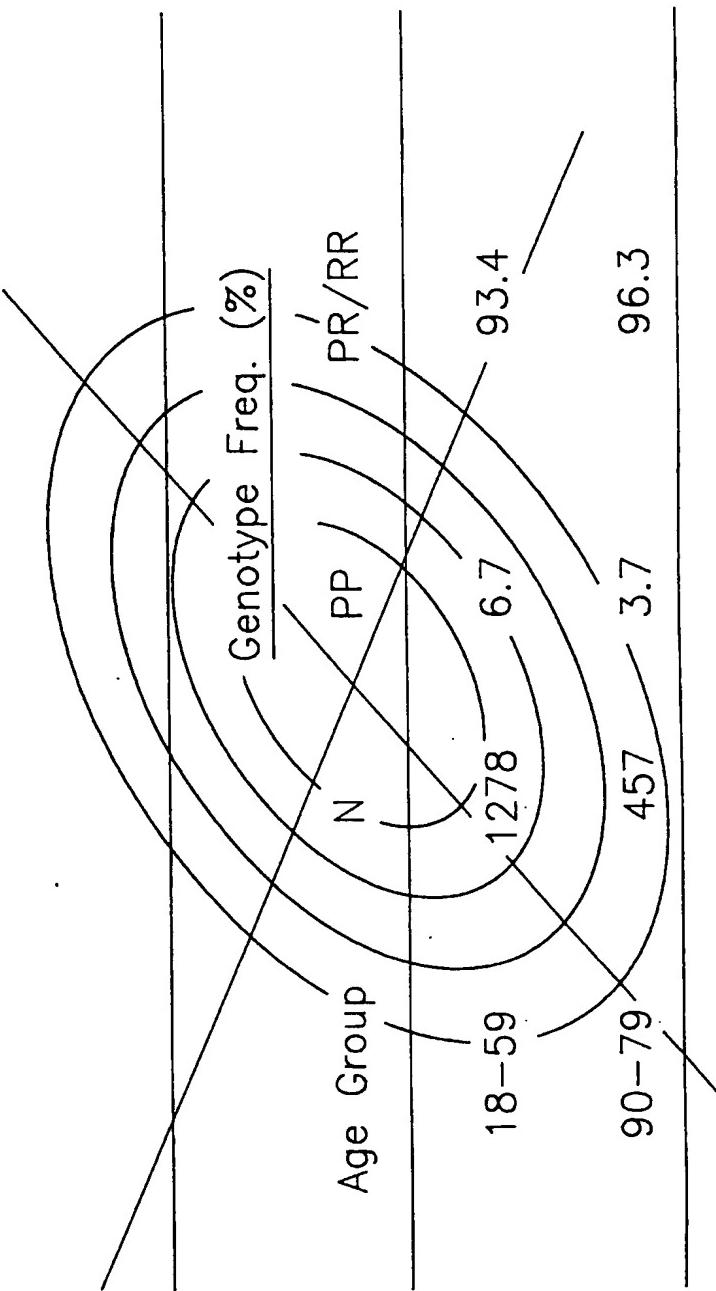


FIG. 6

**FIG. 7A****SUBSTITUTE SHEET (RULE 26)**

12/51

FIG. 7C
P53 PP vs. PR/RR Genotype Distribution
By Age cut point = 59



Sample Size : 1735
 χ^2 : 5.2 (1 d.f.), P = 0.02

14/51

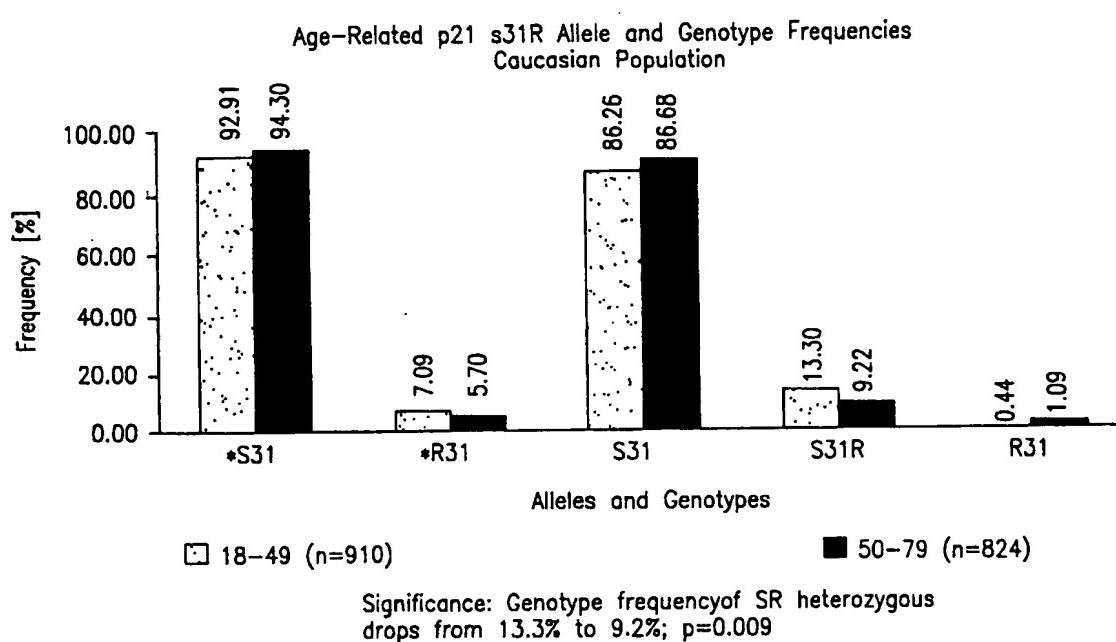


FIG. 8

16 / 51

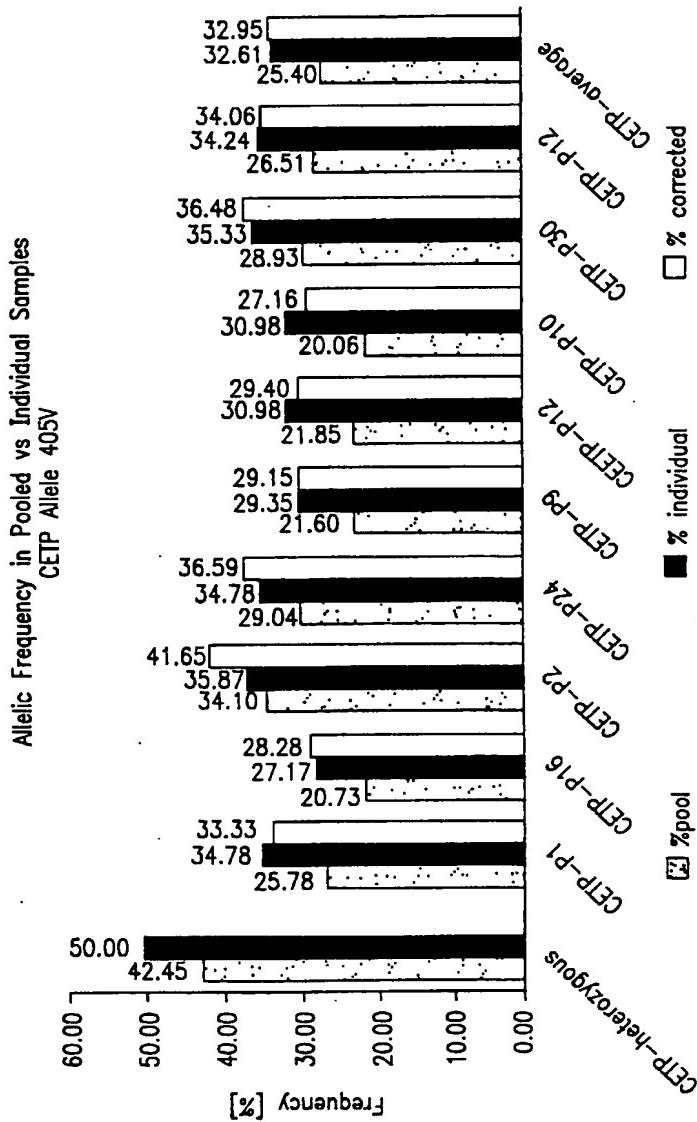


FIG. 10

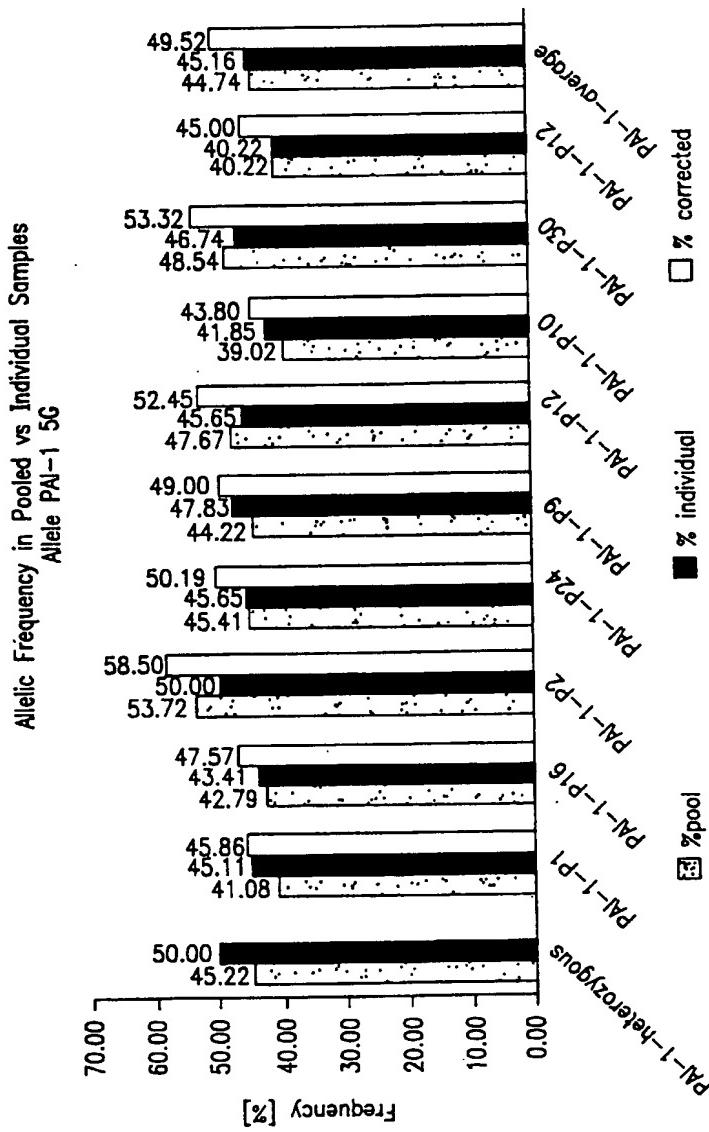


FIG. II

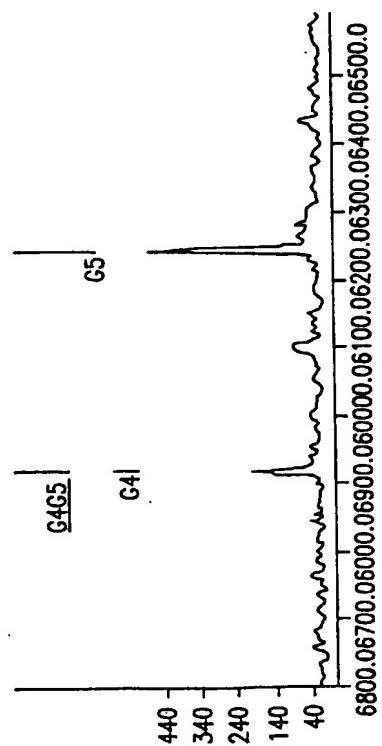
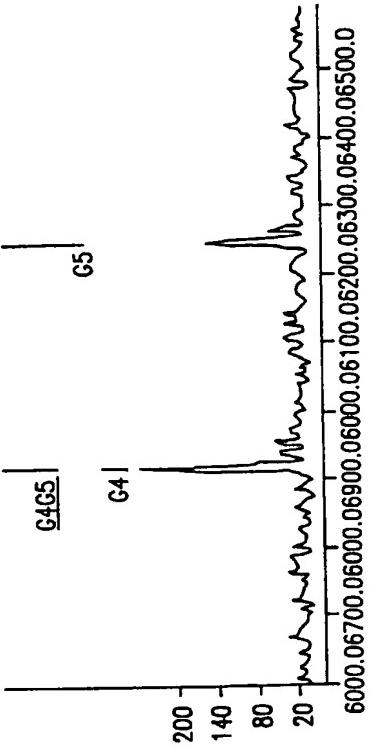
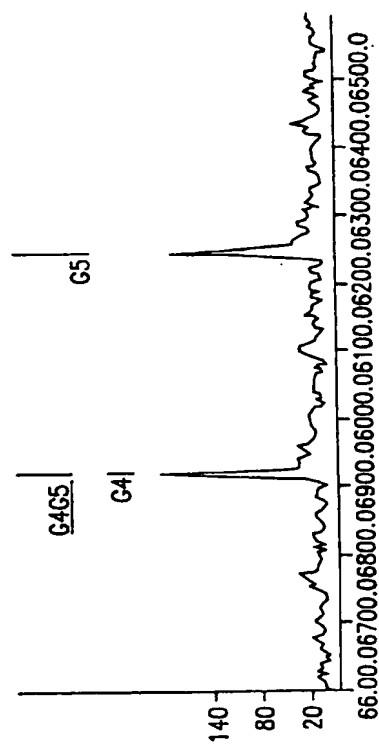


FIG. 12A



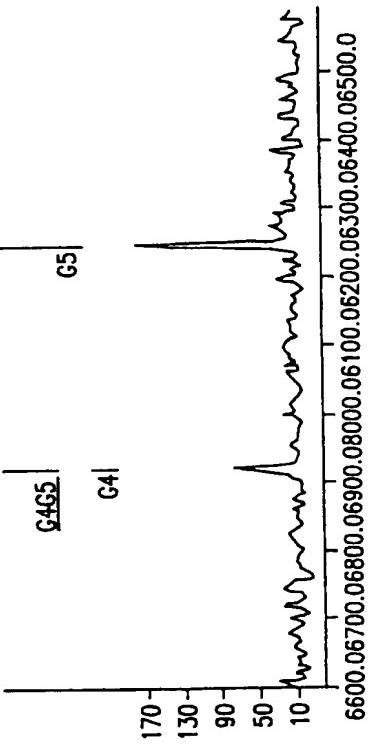
Asian 4G 65.8% 5G 34.2%

FIG. 12B



Caucasian 4G 53.5% 5G 46.5%

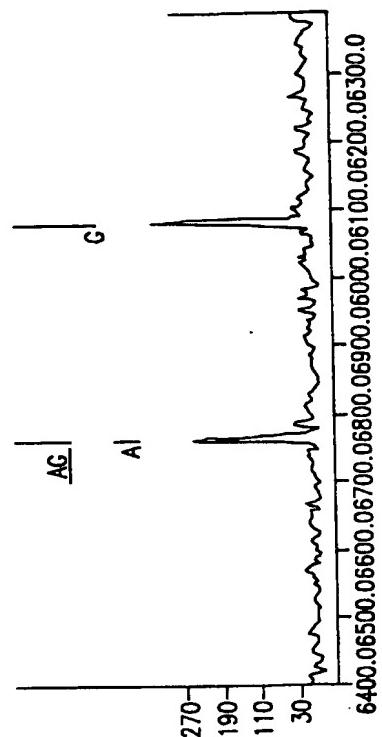
FIG. 12C



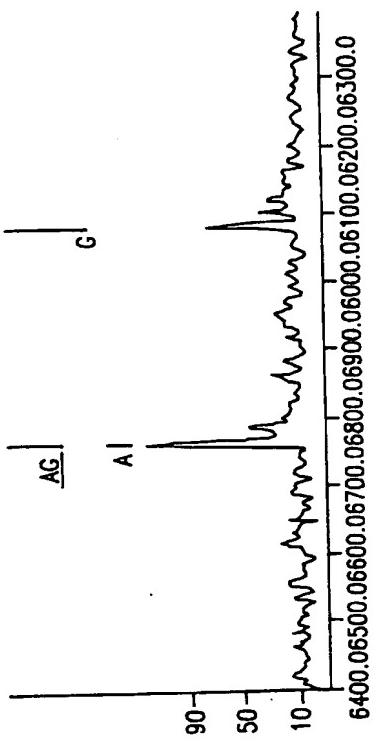
Hispanic 4G 26.7% 5G 73.3%

FIG. 12D

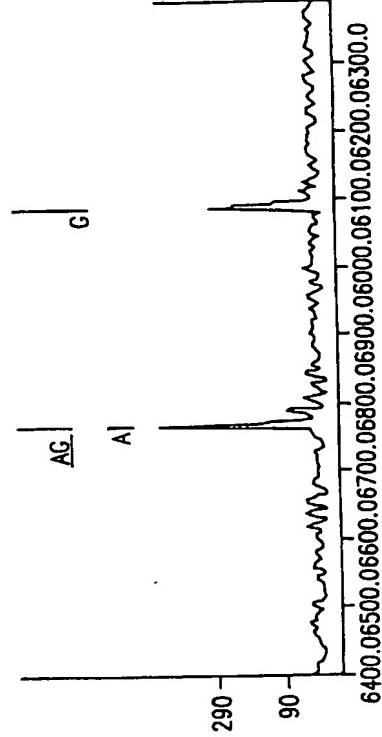
19 / 51

**FIG. 13A**

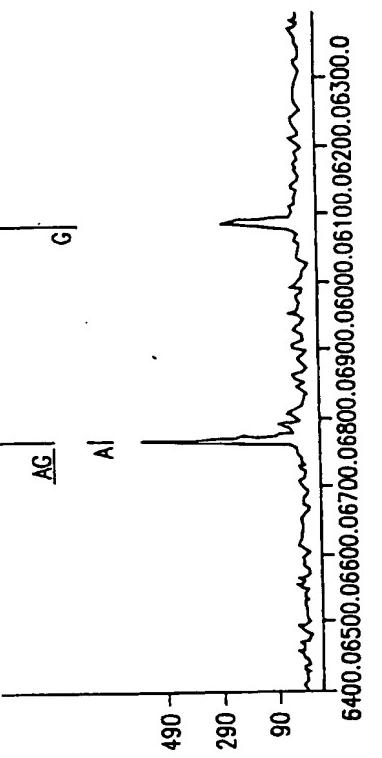
African Ile 43.6% Val 56.4%

**FIG. 13B**

Asian Ile 60.7% Val 39.3%

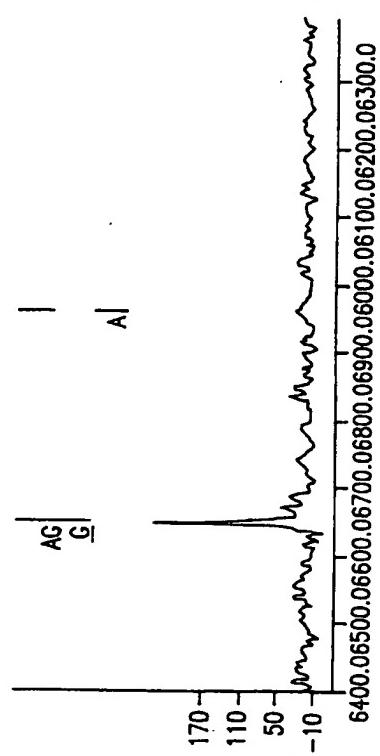
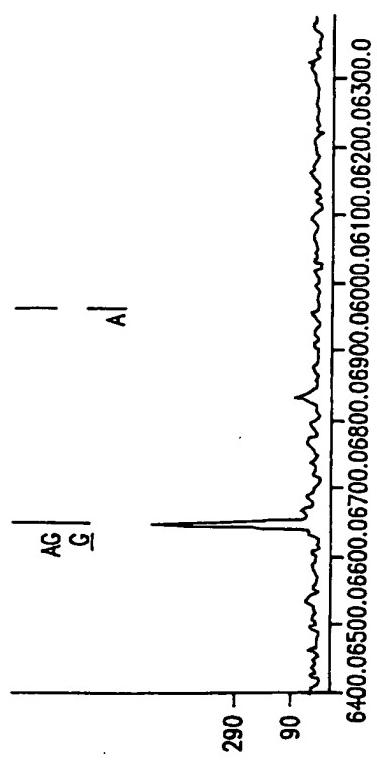
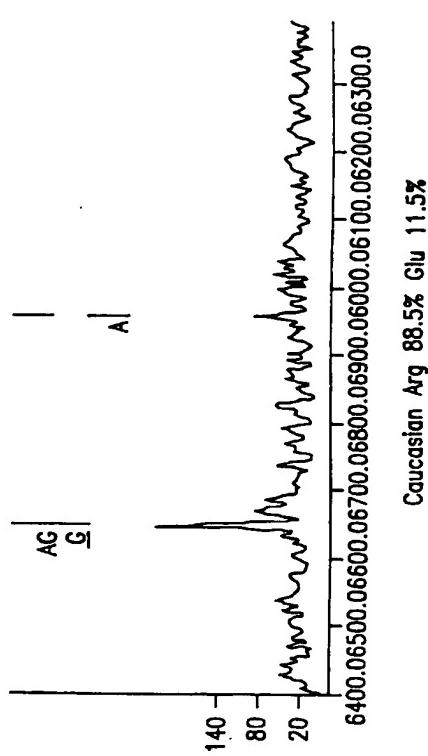
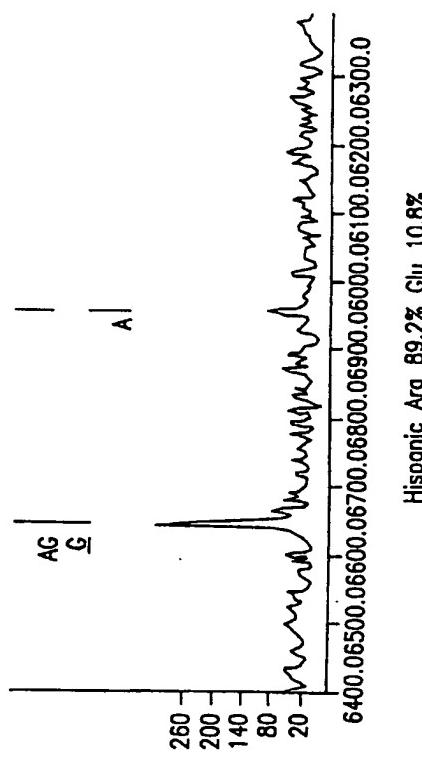
**FIG. 13C**

Caucasian Ile 63.0% Val 37.0%

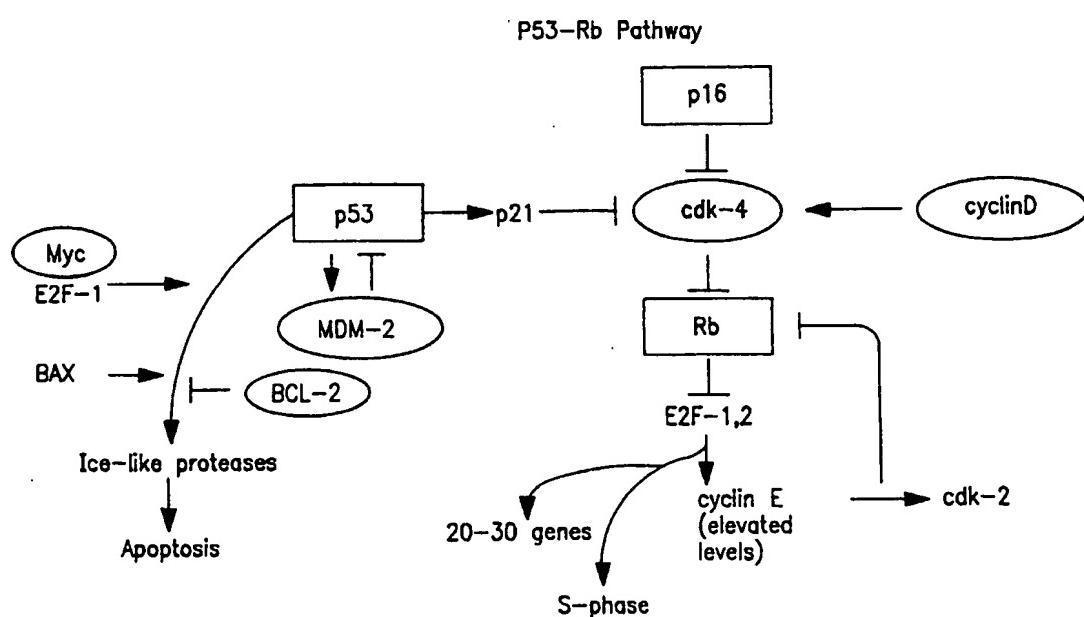
**FIG. 13D**

Hispanic Ile 70.1% Val 29.9%

20/51

**FIG. I4A****FIG. I4B****FIG. I4C****FIG. I4D**

22/ 51

**FIG. 16**

23/51

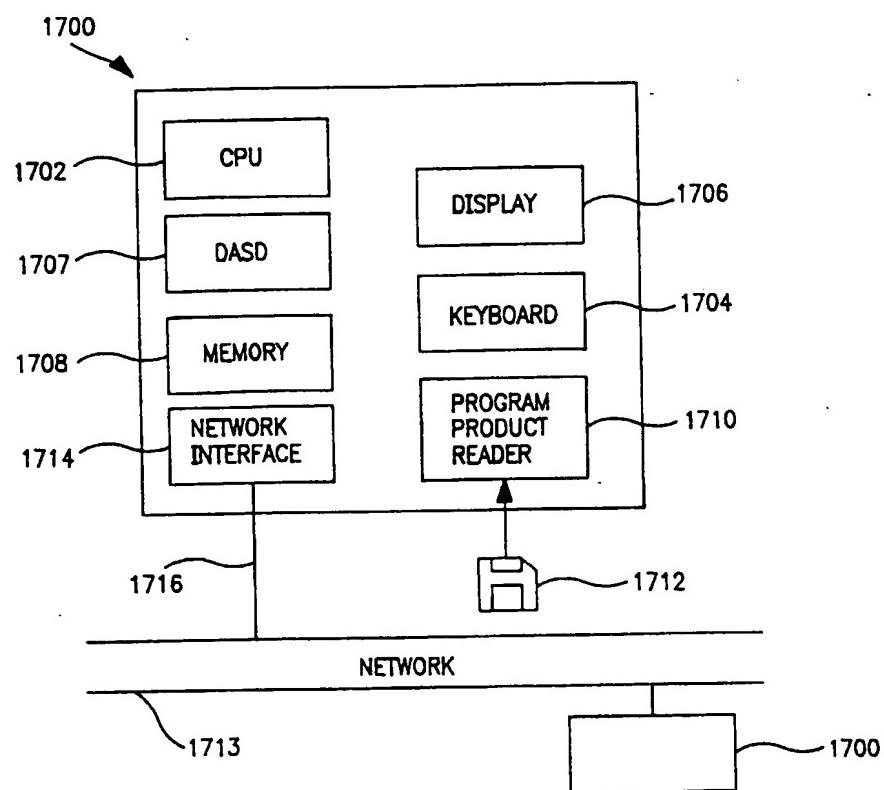


FIG. 17

24/51

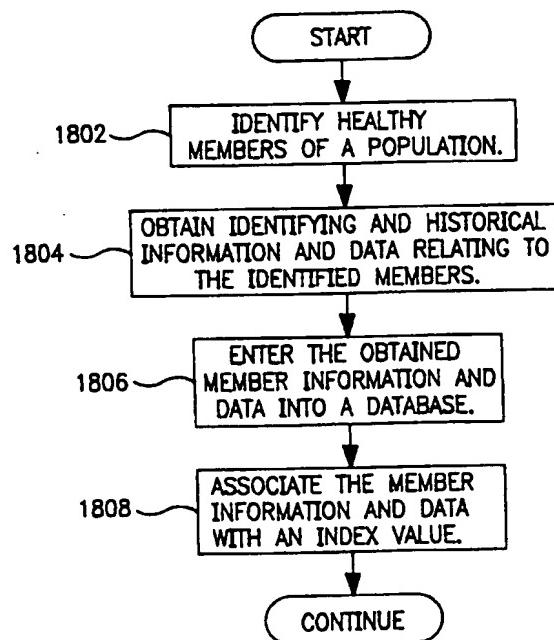


FIG. 18

25/51

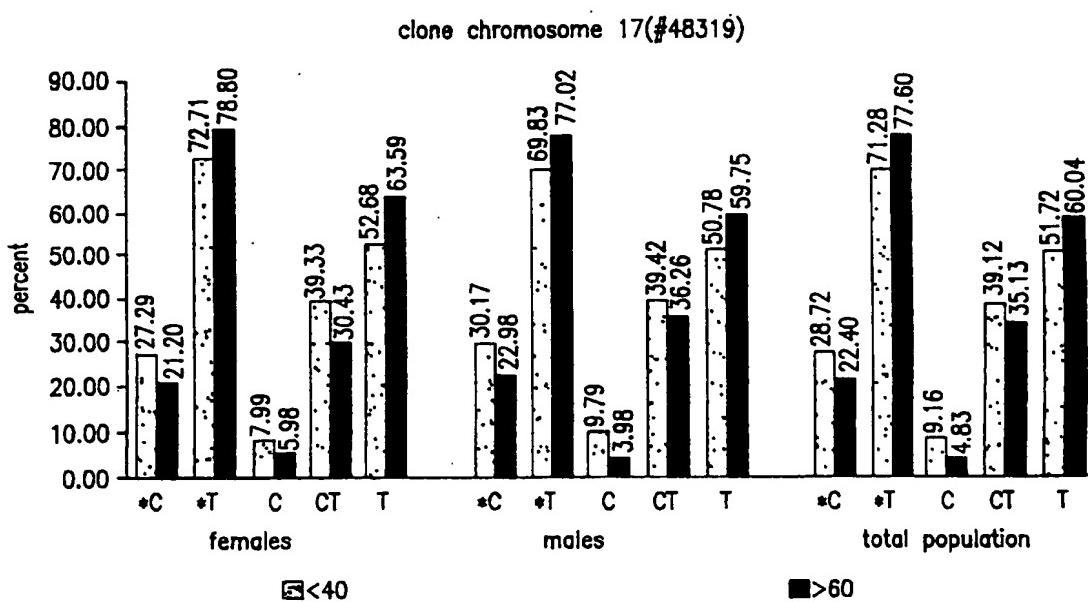


FIG. 19

26/51

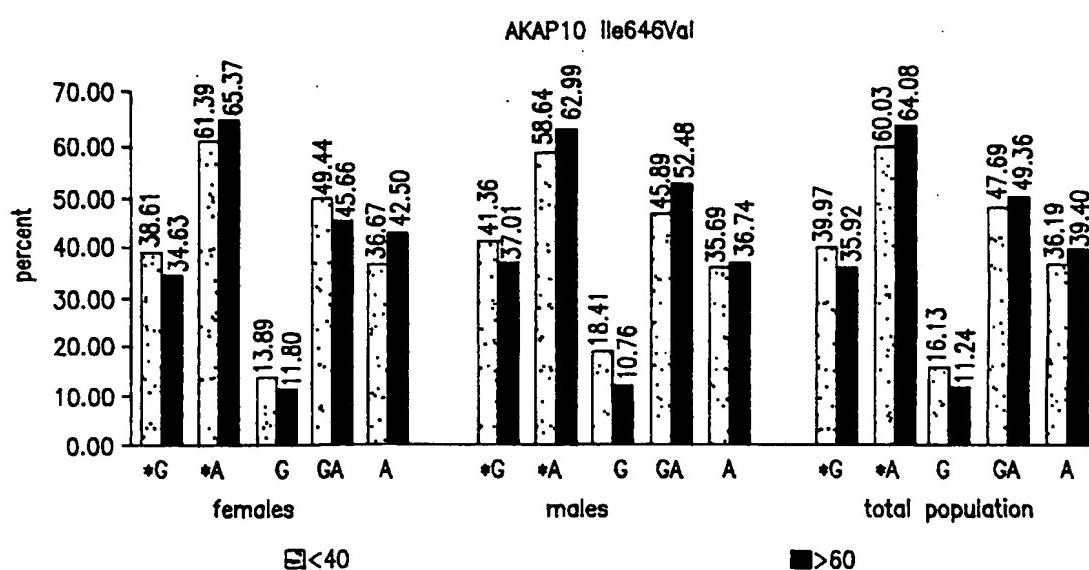


FIG. 20

27/51

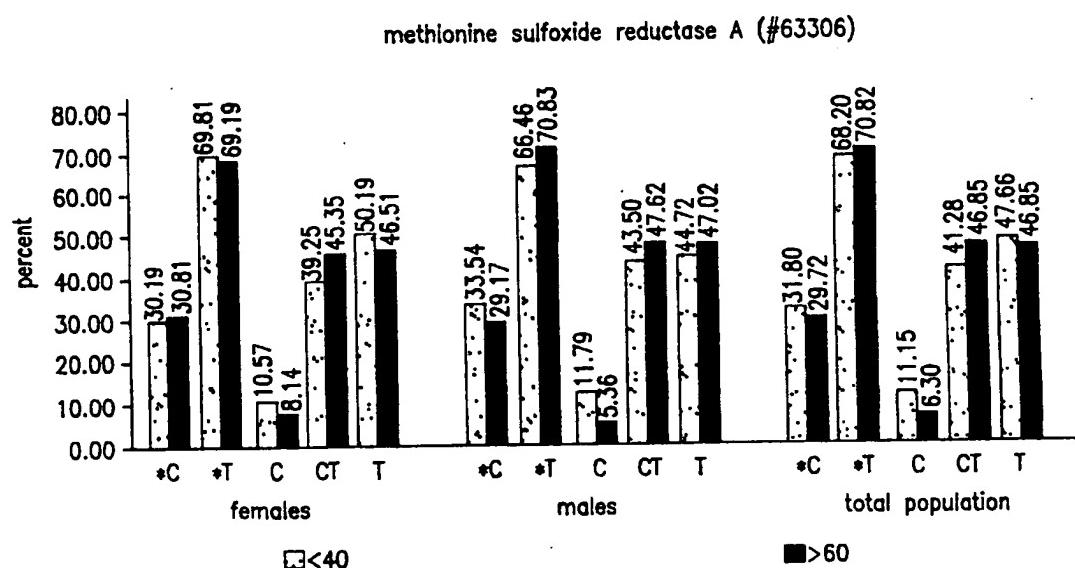


FIG. 2I

28/51

FIG. 22A

29 / 51

Have you ever smoked? <input type="checkbox"/> Yes <input type="checkbox"/> No	Have you been hospitalized in the past 5 years for more than 6 days at a time? <input type="checkbox"/> Yes <input type="checkbox"/> No	1) Weeks: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Acute disorder, including infection and thrombosis <input type="checkbox"/> Chronic Disorder <input type="checkbox"/> Accident <input type="checkbox"/> Other: _____
If yes, for how long?	Years	2) Weeks: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Acute disorder, including infection and thrombosis <input type="checkbox"/> Chronic Disorder <input type="checkbox"/> Accident <input type="checkbox"/> Other: _____
		3) Weeks: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> Acute disorder, including infection and thrombosis <input type="checkbox"/> Chronic Disorder <input type="checkbox"/> Accident <input type="checkbox"/> Other: _____
		For each hospitalization (if not the same) how long did you stay and for what reason?

Have you or has anyone in your immediate family (parents, brothers, sisters, or your children) had the following? Mark all that apply.

Mark all that apply:	You	Mother	Father	Sister	Brother	Child
Heart Disease, including arteriosclerosis	<input type="checkbox"/>					
Stroke	<input type="checkbox"/>					
Hypertension	<input type="checkbox"/>					
Blood clots	<input type="checkbox"/>					
Diabetes, insulin dependent	<input type="checkbox"/>					
Diabetes, not insulin-dependent (diet controlled)	<input type="checkbox"/>					
Cancer:						
Lung & Bronchus	<input type="checkbox"/>					
Breasts	<input type="checkbox"/>					
Prostate	<input type="checkbox"/>					
Colon & Rectum	<input type="checkbox"/>					
Skin	<input type="checkbox"/>					
Lymphoma & Leukemia	<input type="checkbox"/>					
Other, please specify below:						
Alzheimer's Disease	<input type="checkbox"/>					
Epilepsy	<input type="checkbox"/>					
Schizophrenia	<input type="checkbox"/>					
Bipolar disorder (manic depression)	<input type="checkbox"/>					
Major depression	<input type="checkbox"/>					
Chronic Inflammatory or Autoimmune Disease including Multiple Sclerosis and Rheumatoid Arthritis	<input type="checkbox"/>					
Emphysema	<input type="checkbox"/>					
Asthma	<input type="checkbox"/>					
Other, please specify below:						

Do you take prescription drugs on a regular basis?

Yes No

If yes, please specify below:

Have you ever donated blood before? Yes No Additional health information details you would like to provide:

If yes, how many times: Number of Times

**FOR
OFFICE
USE ONLY**

Do you drink any kind of alcoholic beverage?

Never Hardly ever
 Less than 3 times per week 3 or more times per week
 Daily

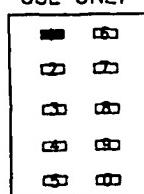


FIG. 22B

SUBSTITUTE SHEET (RULE 26)

30/51

FIG. 22C

SUBSTITUTE SHEET (RULE 26)

32 / 51

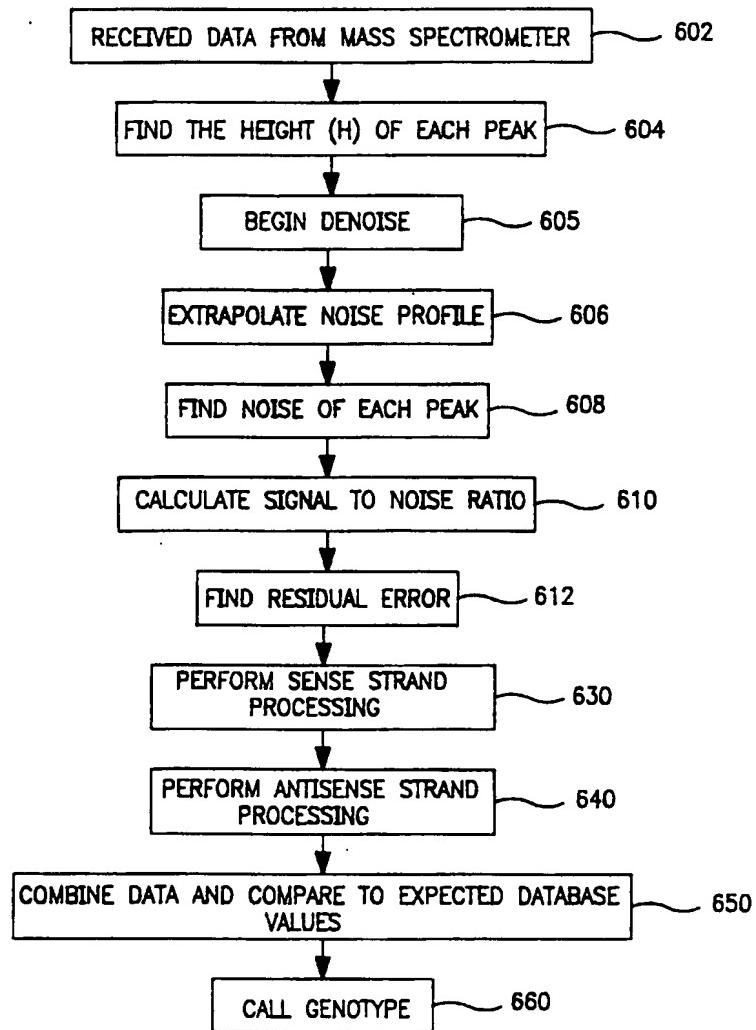


FIG. 23

34 / 51

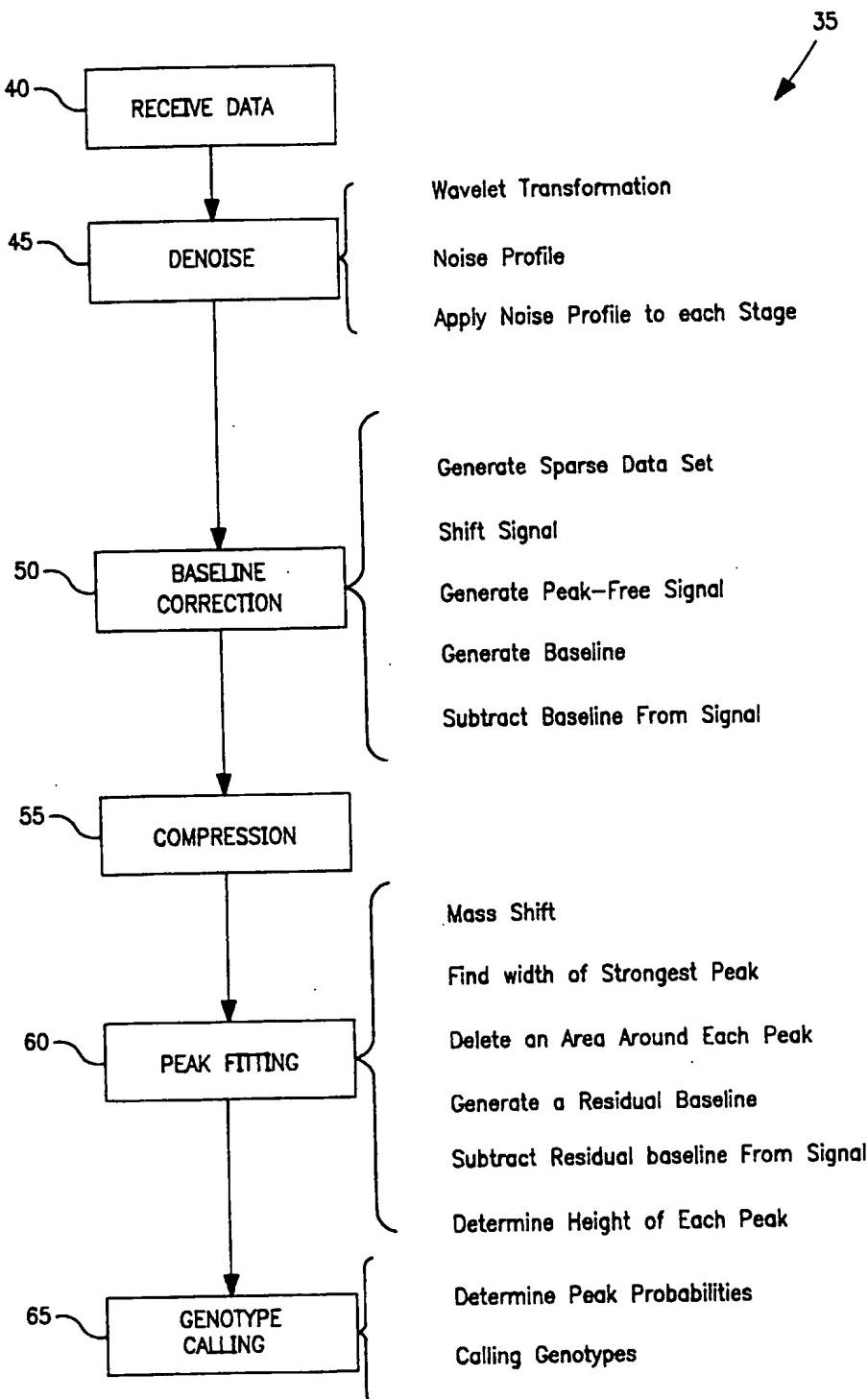


FIG. 25

SUBSTITUTE SHEET (RULE 26)

35/51

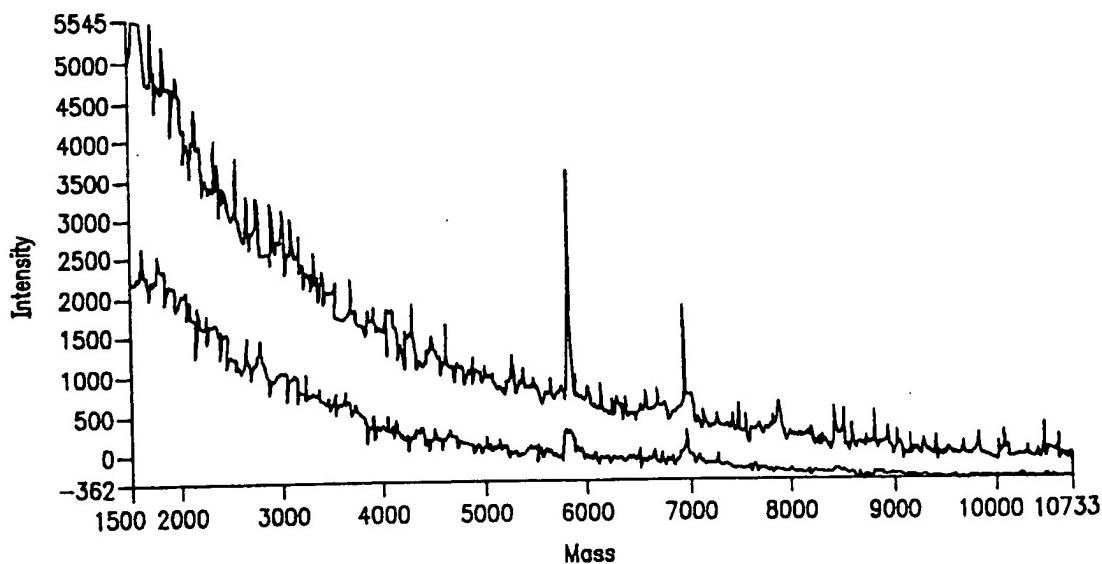


FIG. 26

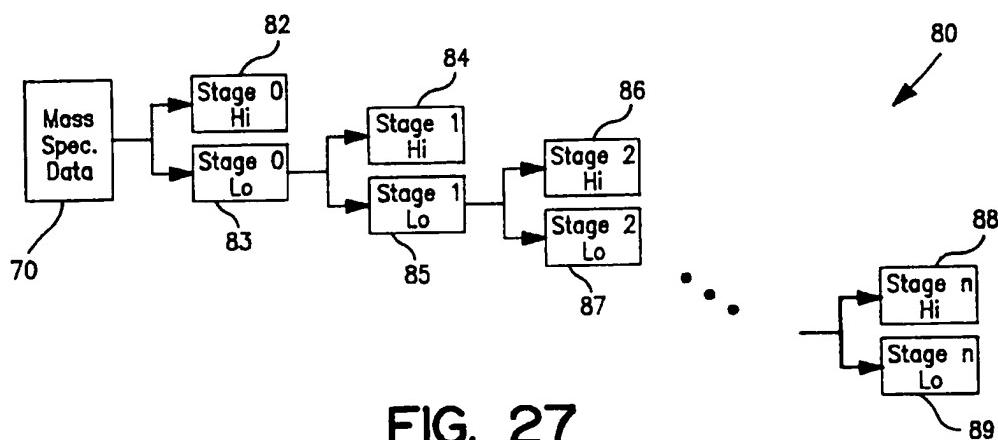


FIG. 27

37/ 51

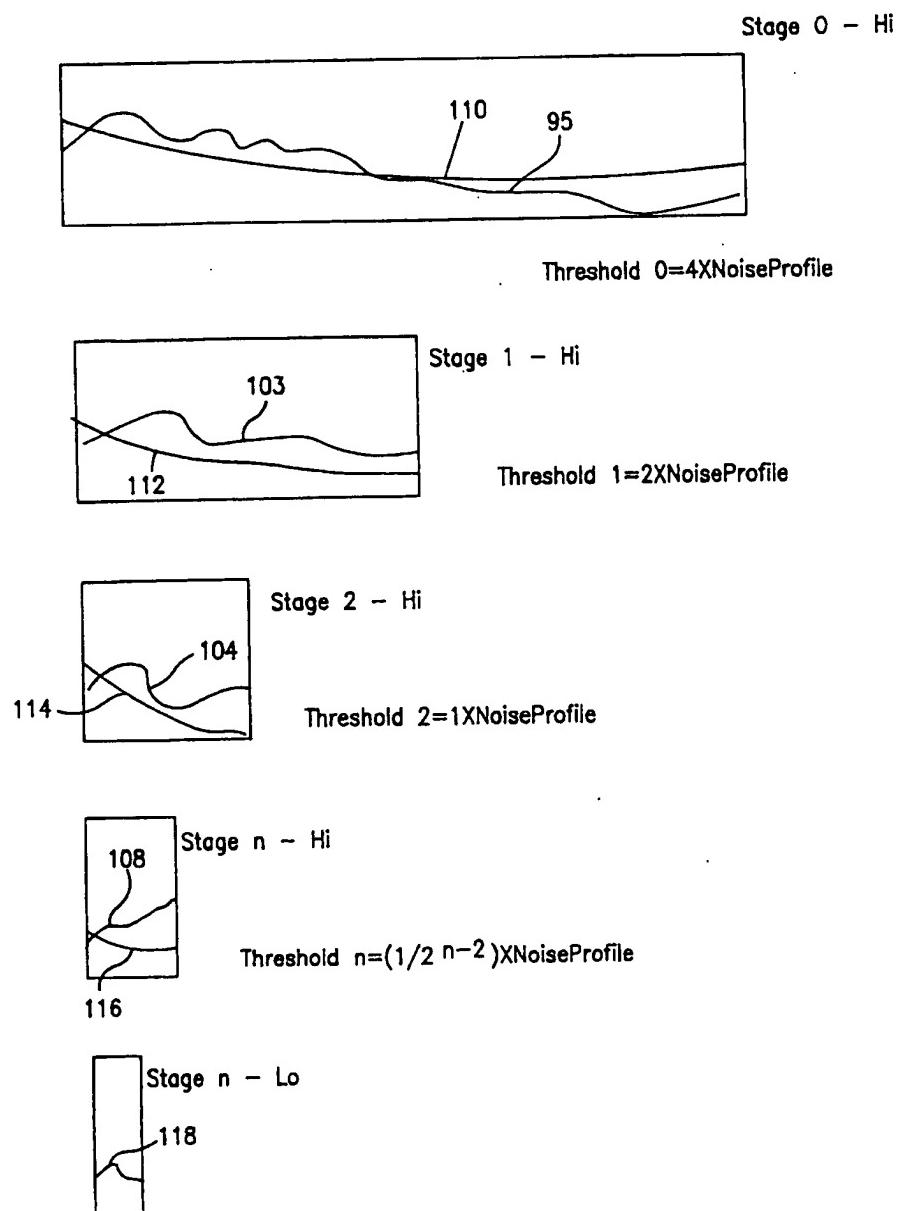


FIG. 3I

38/51

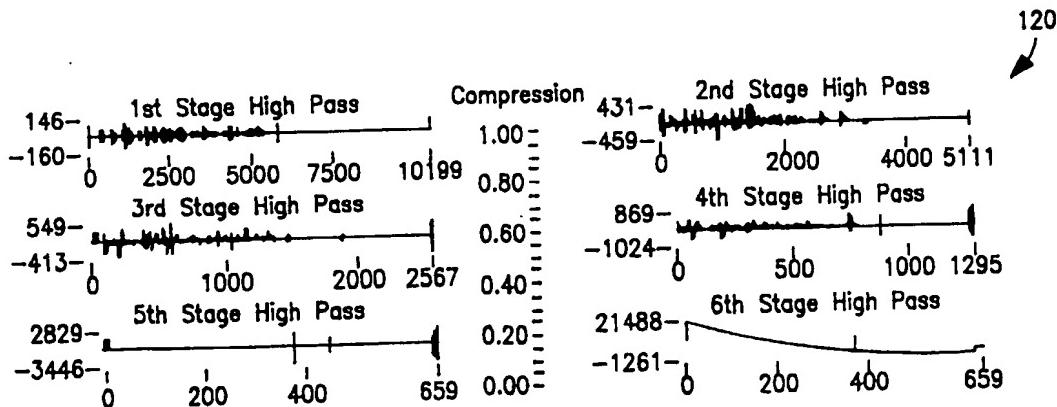


FIG. 32

$$\text{Signal } (t) = \frac{(\text{Start } 0(t) + \text{Start } 1(t) + \text{Start } 2(t) \dots + \text{Start } 23(t))}{24}$$

SHIFT SIGNAL TO ACCOUNT FOR
VARIATIONS DUE TO STARTING POINT

FIG. 33

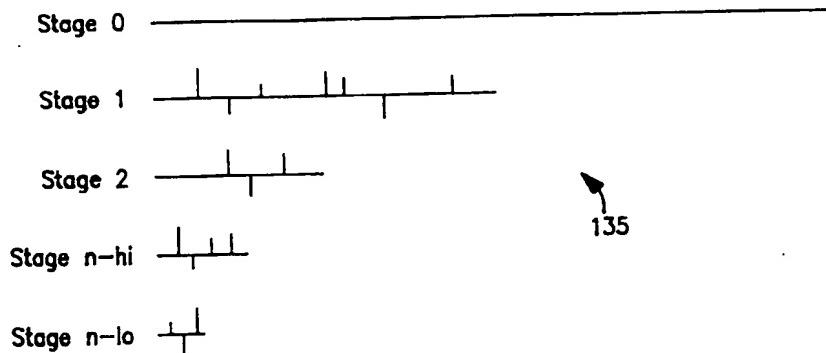


FIG. 34

SUBSTITUTE SHEET (RULE 26)

39/51

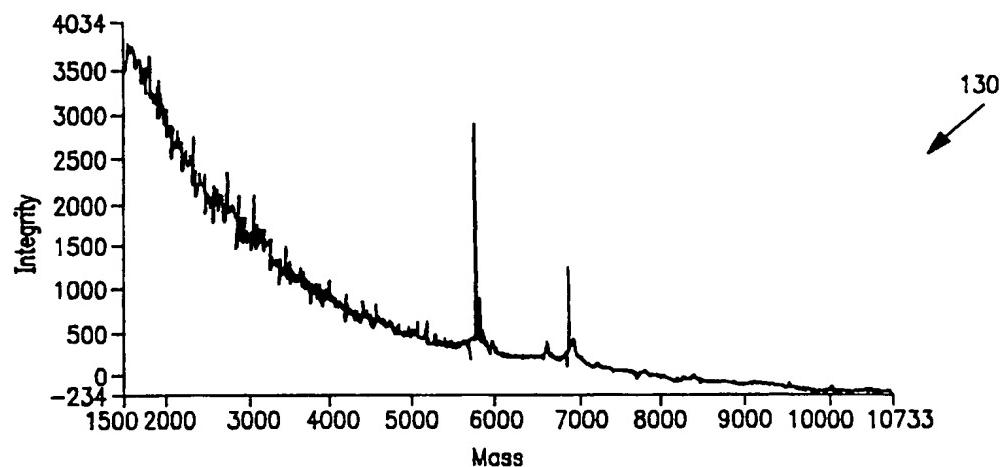


FIG. 35

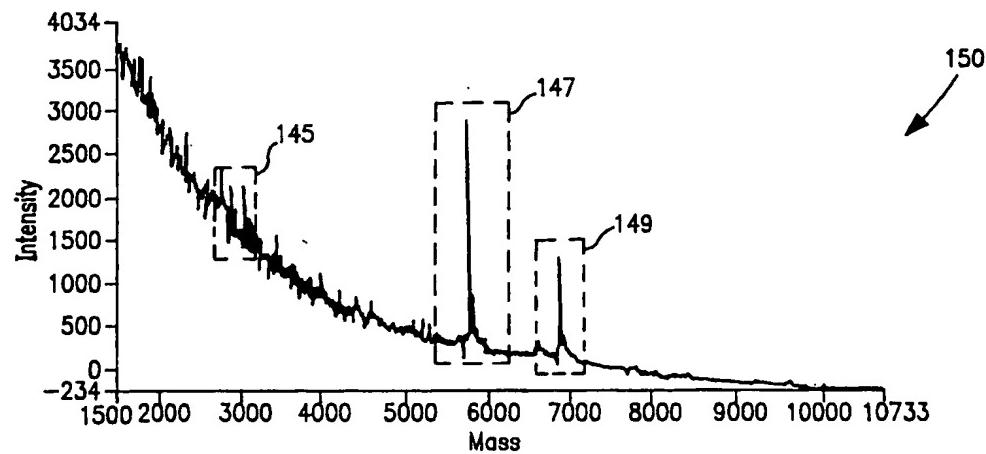


FIG. 13—TAKE A MOVING AVERAGE, REMOVE SECTIONS EXCEEDING A THRESHOLD

FIG. 36

41 / 51

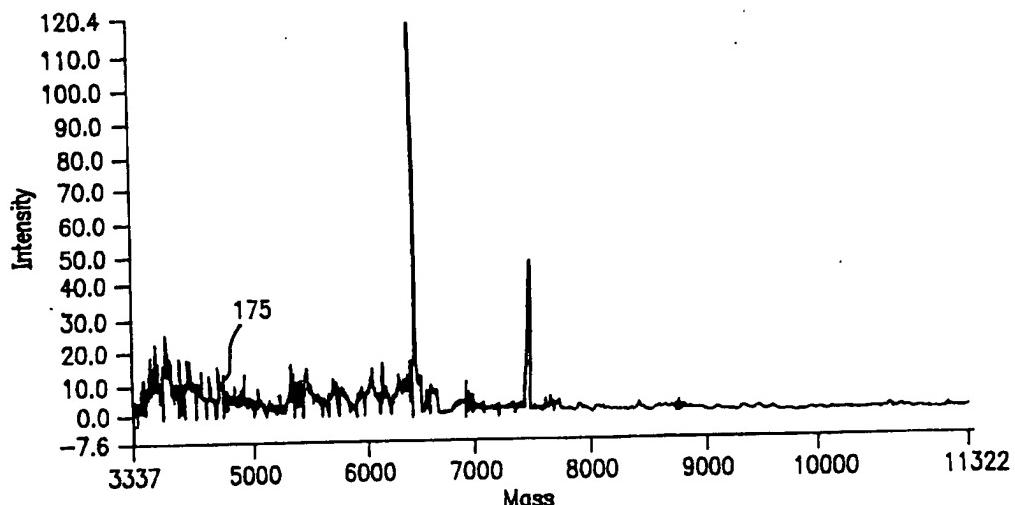


FIG. 40

NON-0 COEFFICIENTS	COEFFICIENTS	INTERMEDIATE	RELATIVE
183	100 150 500 10,050 10,075 11,125 12,100 13,250	25 220 800 890 910 1000 (MAX) 940	100.025 150.220 500.0001 10,050.8 10,075.89 11,125.91 12,100.99999 13,250.94
	.1		100.025 50.220 350.0001 9550.8 25.89 150.91 975.99999 1150.94

FIG. 41

43/51

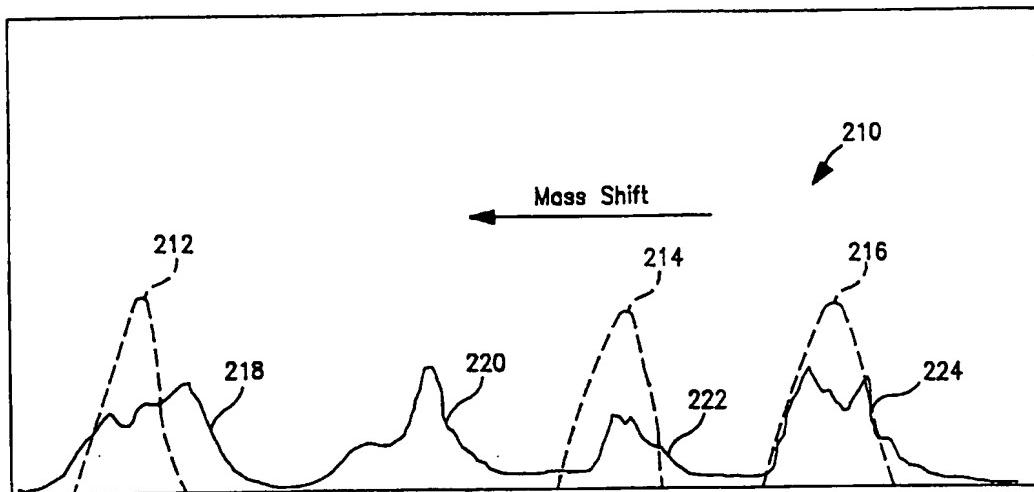


FIG. 43

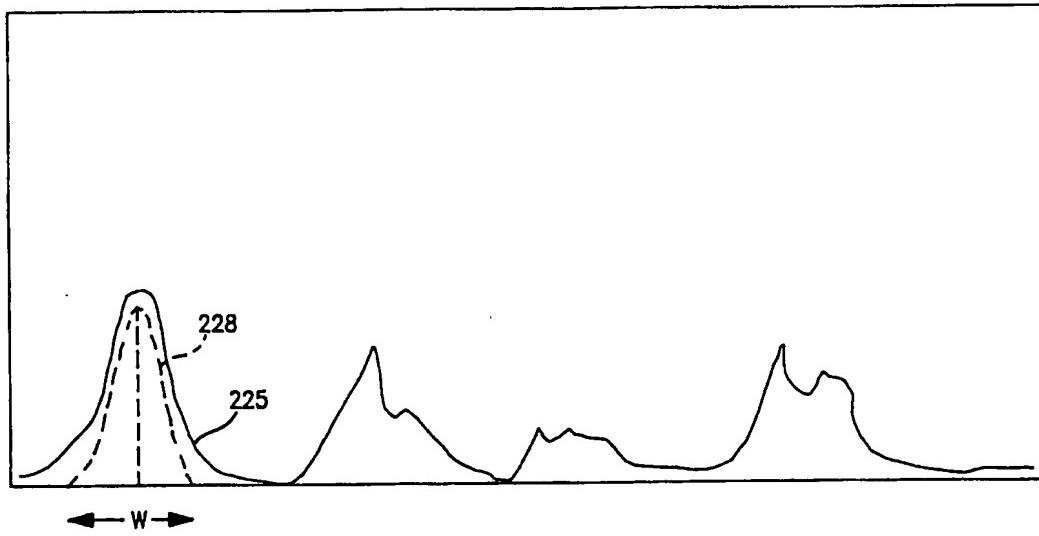


FIG. 44

44/51

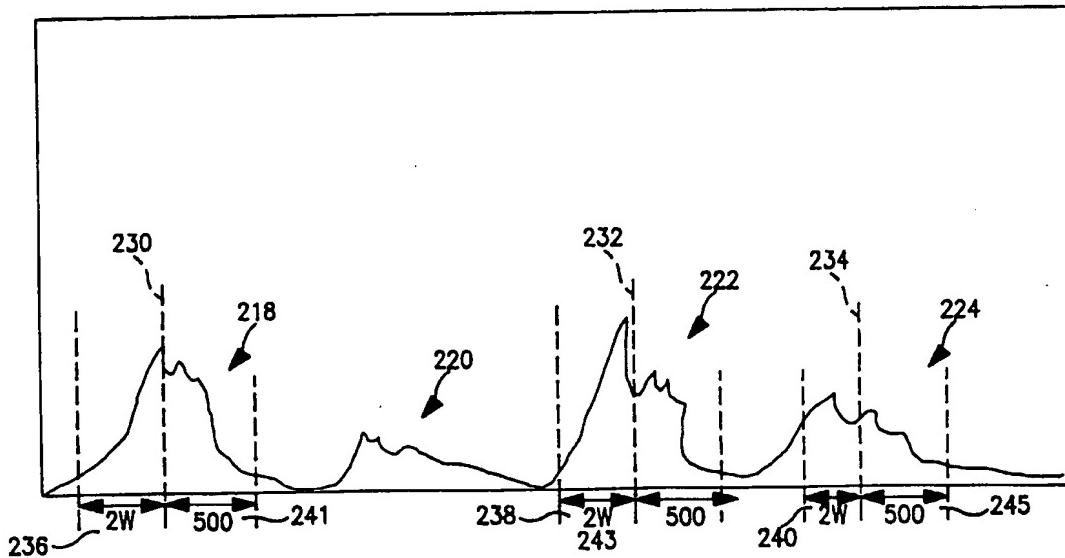


FIG. 45

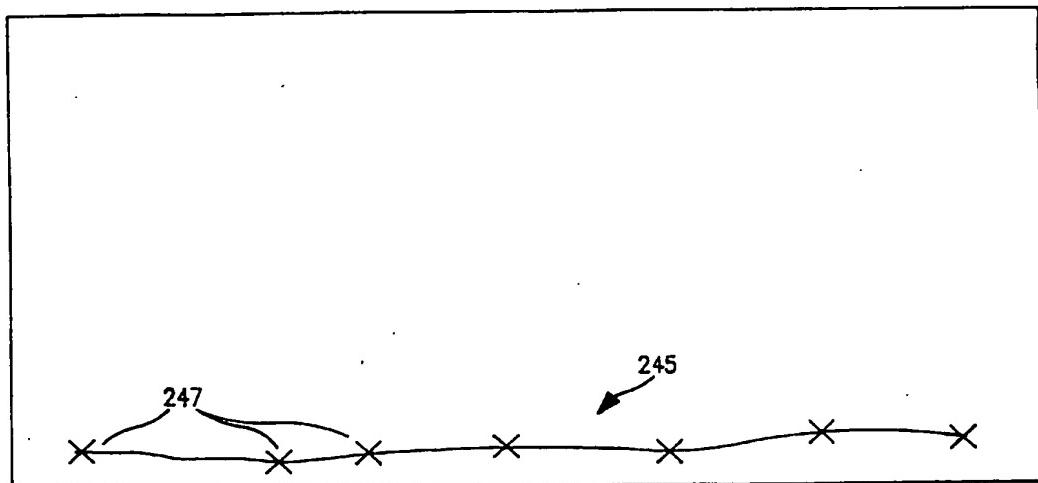


FIG. 46

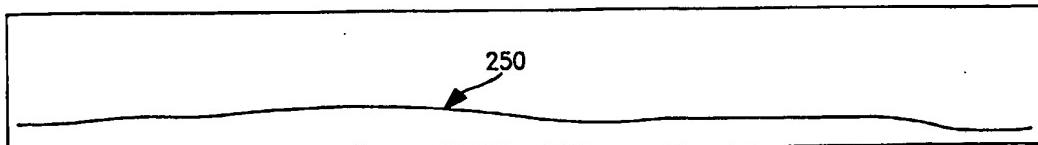


FIG. 47

SUBSTITUTE SHEET (RULE 26)

FIG. 51

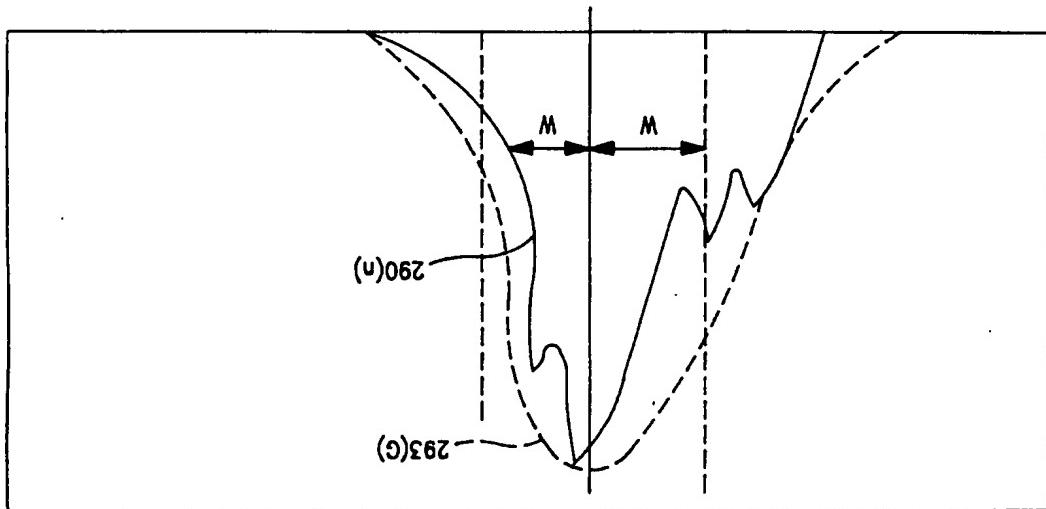
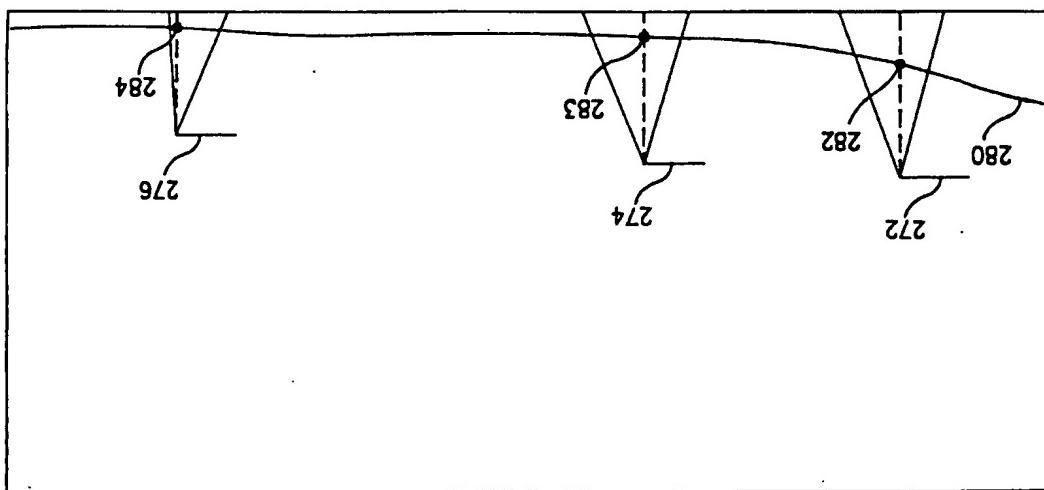


FIG. 50



46/51

SUBSTITUTE SHEET (RULE 26)

FIG. 53

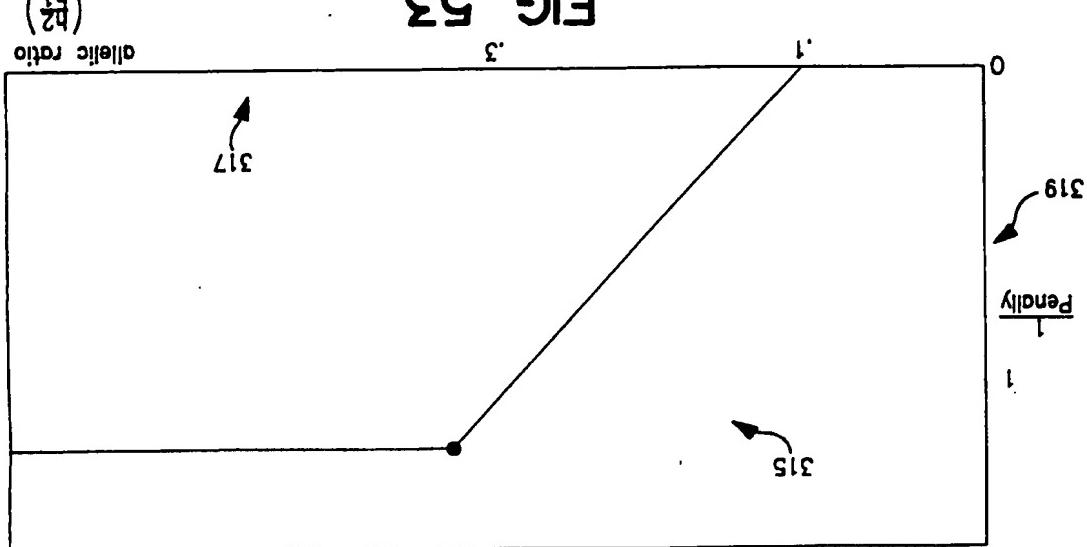
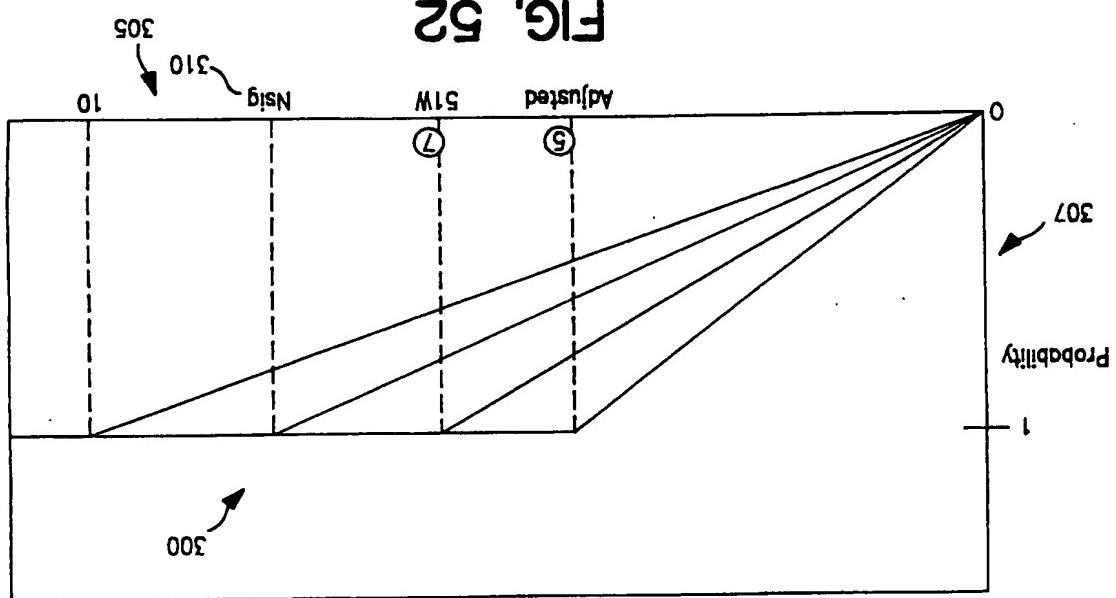


FIG. 52



47/51

SUBSTITUTE SHEET (RULE 26)

FIG. 55

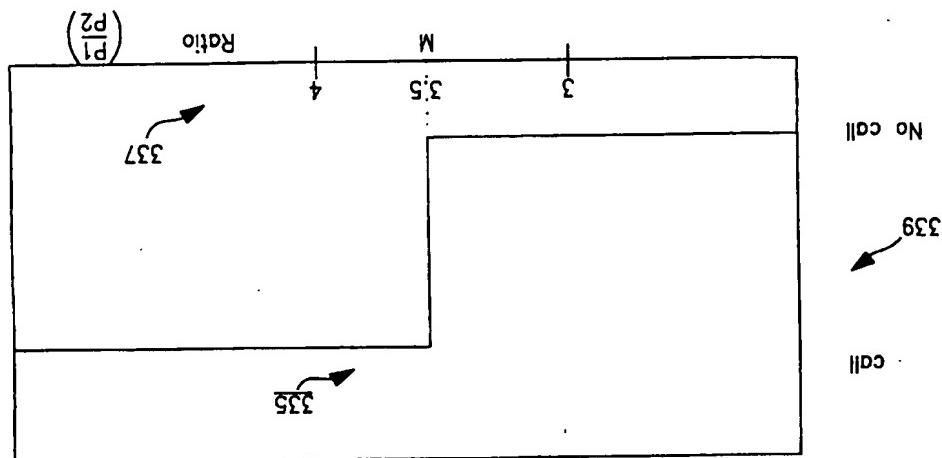
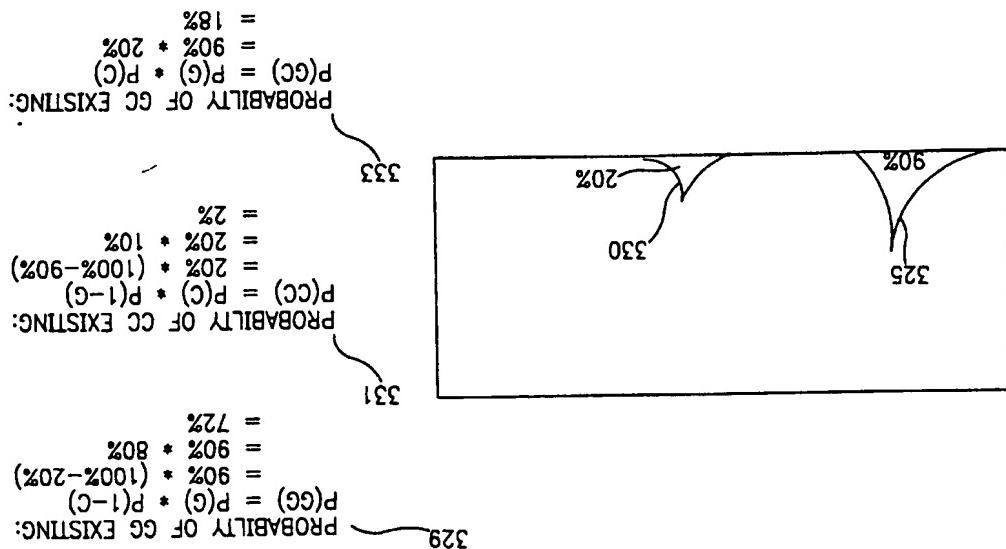
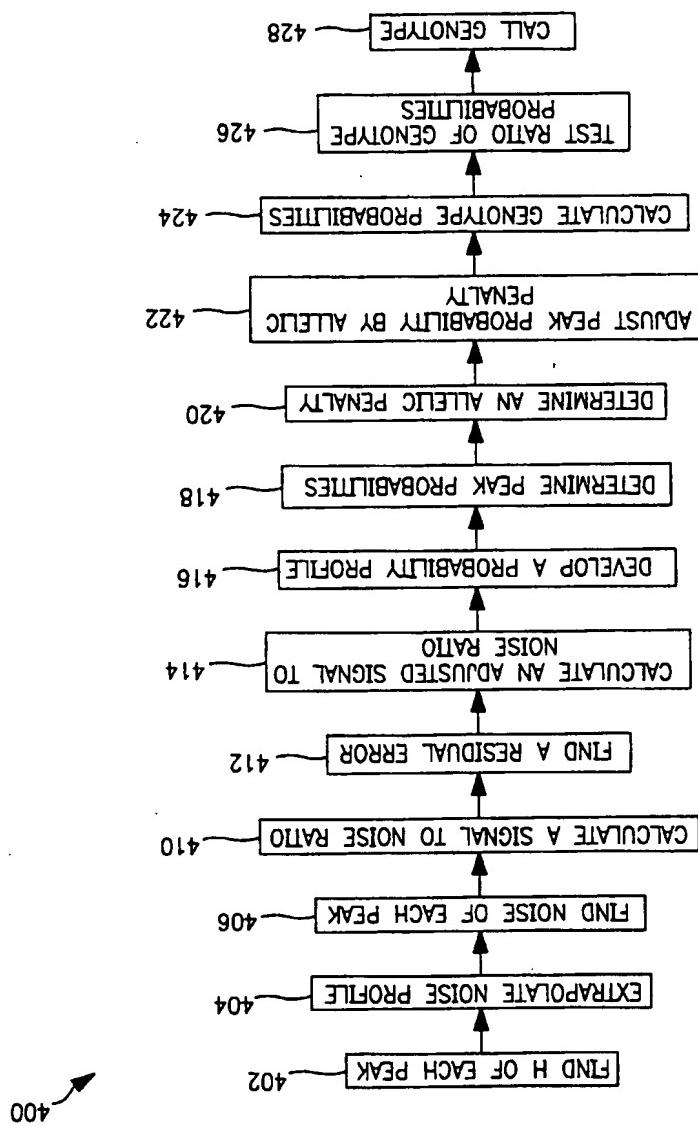


FIG. 54



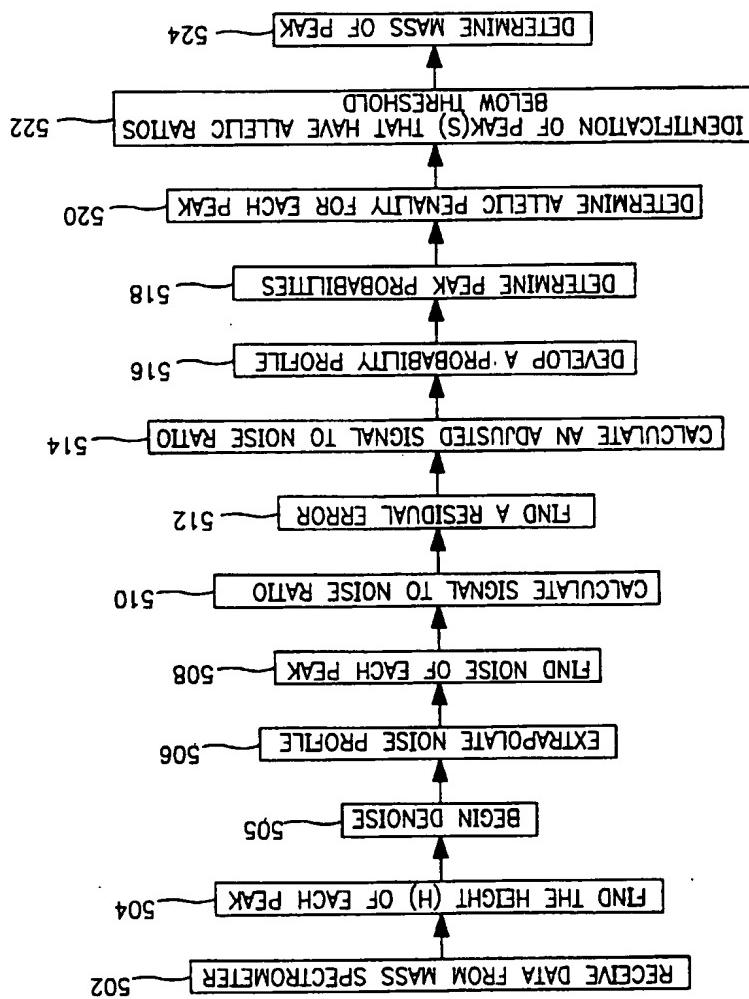
SUBSTITUTE SHEET (RULE 26)

FIG. 56



49/51

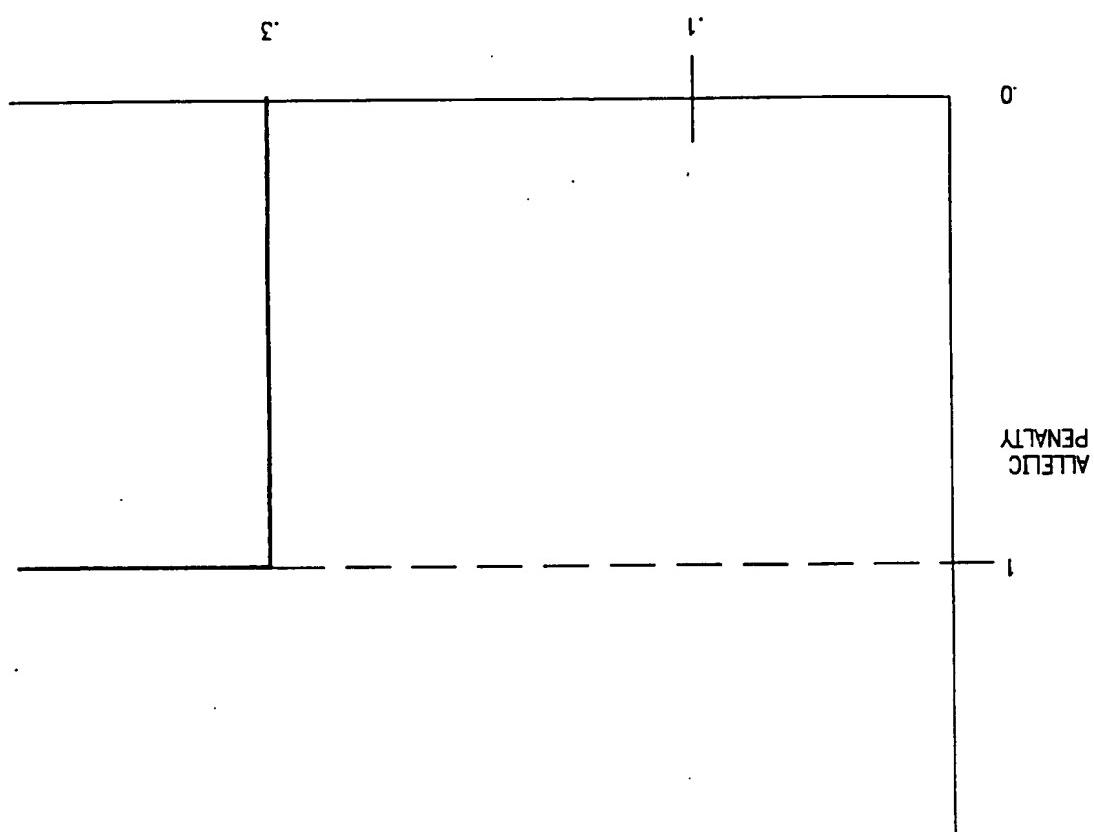
FIG. 57



SUBSTITUTE SHEET (RULE 26)

FIG. 58

RATIO OF AREA UNDER PEAK



51/51

SEQUENCE LISTING

WO 01/27857

PCT/US00/28413

2/122

<213> Artificial Sequence

<220>
<223> Probe

<400> 4
gccagaggct gctcccc 17

<210> 5
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 5
gccagaggct gctcccc 17

<210> 6
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 6
gccagaggct gctccccc 19

<210> 7
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 7
gccagaggct gctccccc 18

<210> 8
<211> 161
<212> DNA
<213> Homo Sapien

<400> 8
gtccgtcaga acccatgcgg cagcaaggcc tgccgcccccc ttttcggccc agtggacagg 60
gaggagctga gccgcgactg tggatgcgcta atggcggttgc gcatccagga ggcccggttag 120
cgatggaaact tcgactttgt caccgagaca ccactggagg g 161

<210> 9
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 9

3/122

cccaagtacg acgttgtaaa acggtccgtc agaacccatg cgg 43
<210> 10
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 10
agcggtataac aatttcacac aggctccagt ggtgtctcgg tgac 44
<210> 11
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 11
cagcgagcag ctgag 15
<210> 12
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 12
cagcgagcag ctgag 15
<210> 13
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 13
cagcgagcag ctgagc 16
<210> 14
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 14
cagcgagcag ctgagac 17
<210> 15
<211> 205
<212> DNA

4/122

<213> Homo Sapien

<400> 15
gcgcgtccatt catctttca tcgactctct gttgaatgaa gaaaatccaa gtaaggccta 60
caggtgcagt tccaaggaag cctttgagaa agggctctgc ttgagttgt aaaaagaaccg 120
ctgcaacaat ctgggctatg agatcaataa agtcagagcc aaaagaagca gaaaaatgt 180
cctgaagact cgttctcaga tgccc 205

<210> 16

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primers

<400> 16
cccagtacg acgttgtaaa acggcgctcc attcatctct tc 42

<210> 17

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer

<400> 17
agcggataac aatttcacac agggggcatac tgagaacgag tc 42

<210> 18

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Oligonucleotide Primer

<400> 18
caatctggc tatgagatca 20

<210> 19

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Probe

<400> 19
caatctggc tatgagatca 20

<210> 20

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Probe

5/122

<400> 20
caatctgggc tatgagatca a 21

<210> 21
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 21
caatctgggc tatgagatca gt 20

<210> 22
<211> 60
<212> DNA
<213> Homo Sapien

<220>
<223> Probe

<400> 22
gtgccggcta ctcggatggc agcaaggact cctgcaaggg ggacagtgga ggccccacatg 60

<210> 23
<211> 60
<212> DNA
<213> Homo sapien

<400> 23
ccacccacta ccggggcacg tggtacctga cgggcacatcg cagctggggc cagggctgcg 60

<210> 24
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 24
cccagtccacg acgttgtaaa acgatggcag caaggactcc tg 42

<210> 25
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 25
cacatgccac ccactacc 18

<210> 26
<211> 43
<212> DNA
<213> Artificial Sequence

6/122

<220>
<223> Oligonucleotide primer

<400> 26
agcggataac aatttcacac aggtgacgat gcccgtcagg tac 43

<210> 27
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 27
atgccaccca ctacc 15

<210> 28
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 28
cacatgccac ccactaccg 19

<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 29
cacatgccac ccactaccag 20

<210> 30
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Probe

<400> 30
agcggataac aatttcacac agg 23

<210> 31
<211> 2363
<212> DNA
<213> Homo Sapien

<220>
<221> CDS
<222> (138) ... (2126)
<223> AKAP-10

9/122

Ser Leu Gln Ala Thr His Pro Leu Gly Phe Asp Asp Val Val Arg Leu 445 450 455	
gaa att gaa tcc aat atc tgc agg gaa ggt ggg cca ctc ccc aac tgt Glu Ile Glu Ser Asn Ile Cys Arg Glu Gly Gly Pro Leu Pro Asn Cys 460 465 470 475	1562
ttc aca act cca tta cgt cag gcc tgg aca acc atg gag aag gtc ttt Phe Thr Thr Pro Leu Arg Gln Ala Trp Thr Thr Met Glu Lys Val Phe 480 485 490	1610
ttg cct ggc ttt ctg tcc agc aat ctt tat tat aaa tat ttg aat gat Leu Pro Gly Phe Leu Ser Ser Asn Leu Tyr Tyr Lys Tyr Leu Asn Asp 495 500 505	1658
ctc atc cat tcg gtt cga gga gat gaa ttt ctg ggc ggg aac gtg tcg Leu Ile His Ser Val Arg Gly Asp Glu Phe Leu Gly Gly Asn Val Ser 510 515 520	1706
ccg act gct cct ggc tct gtt ggc cct cct gat gag tct cac cca ggg Pro Thr Ala Pro Gly Ser Val Gly Pro Pro Asp Glu Ser His Pro Gly 525 530 535	1754
agt tct gac agc tct gcg tct cag tcc agt gtg aaa aaa gcc agt att Ser Ser Asp Ser Ser Ala Ser Gln Ser Ser Val Lys Lys Ala Ser Ile 540 545 550 555	1802
aaa ata ctg aaa aat ttt gat gaa gcg ata att gtg gat gcg gca agt Lys Ile Leu Lys Asn Phe Asp Glu Ala Ile Ile Val Asp Ala Ala Ser 560 565 570	1850
ctg gat cca gaa tct tta tat caa cgg aca tat gcc ggg aag atg aca Leu Asp Pro Glu Ser Leu Tyr Gln Arg Thr Tyr Ala Gly Lys Met Thr 575 580 585	1898
ttt gga aga gtg agt gac ttg ggg caa ttc atc cgg gaa tct gag cct Phe Gly Arg Val Ser Asp Leu Gly Gln Phe Ile Arg Glu Ser Glu Pro 590 595 600	1946
gaa cct gat gta agg aaa tca aaa gga tcc atg ttc tca caa gct atg Glu Pro Asp Val Arg Lys Ser Lys Gly Ser Met Phe Ser Gln Ala Met 605 610 615	1994
aag aaa tgg gtg caa gga aat act gat gag gcc cag gaa gag cta gct Lys Lys Trp Val Gln Gly Asn Thr Asp Glu Ala Gln Glu Glu Leu Ala 620 625 630 635	2042
tgg aag att gct aaa atg ata gtc agt gac att atg cag cag gct cag Trp Lys Ile Ala Lys Met Ile Val Ser Asp Ile Met Gln Gln Ala Gln 640 645 650	2090
tat gat caa ccg tta gag aaa tct aca aag tta tga ctcaaaaactt Tyr Asp Gln Pro Leu Glu Lys Ser Thr Lys Leu *	2136
655 660	
gagataaaagg aaatctgctt gtgaaaaata agagaacttt tttcccttgg ttggattctt caacacagcc aatgaaaaca gcactatatt tctgatctgt cactgttgg tccagggaga gaatggggag acaa c cttag gacttccacc ctaatgcagt tacctgttagg gcataattgg atggcacatg atgttcaca cagtgaggag tctttaaagg ttaccaa	2196 2256 2316 2363

10/122

<210> 32
<211> 662
<212> PRT
<213> Homo Sapien

<400> 32
Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser Pro Arg Thr Leu Arg
1 5 10 15
Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg Arg Lys Val Lys Gly
20 25 30
Lys Glu Gln Glu Lys Thr Ser Asp Val Lys Ser Ile Lys Ala Ser Ile
35 40 45
Ser Val His Ser Pro Gln Lys Ser Thr Lys Asn His Ala Leu Leu Glu
50 55 60
Ala Ala Gly Pro Ser His Val Ala Ile Asn Ala Ile Ser Ala Asn Met
65 70 75 80
Asp Ser Phe Ser Ser Arg Thr Ala Thr Leu Lys Lys Gln Pro Ser
85 90 95
His Met Glu Ala Ala His Phe Gly Asp Leu Gly Arg Ser Cys Leu Asp
100 105 110
Tyr Gln Thr Gln Glu Thr Lys Ser Ser Leu Ser Lys Thr Leu Glu Gln
115 120 125
Val Leu His Asp Thr Ile Val Leu Pro Tyr Phe Ile Gln Phe Met Glu
130 135 140
Leu Arg Arg Met Glu His Leu Val Lys Phe Trp Leu Glu Ala Glu Ser
145 150 155 160
Phe His Ser Thr Thr Trp Ser Arg Ile Arg Ala His Ser Leu Asn Thr
165 170 175
Met Lys Gln Ser Ser Leu Ala Glu Pro Val Ser Pro Ser Lys Lys His
180 185 190
Glu Thr Thr Ala Ser Phe Leu Thr Asp Ser Leu Asp Lys Arg Leu Glu
195 200 205
Asp Ser Gly Ser Ala Gln Leu Phe Met Thr His Ser Glu Gly Ile Asp
210 215 220
Leu Asn Asn Arg Thr Asn Ser Thr Gln Asn His Leu Leu Ser Gln
225 230 235 240
Glu Cys Asp Ser Ala His Ser Leu Arg Leu Glu Met Ala Arg Ala Gly
245 250 255
Thr His Gln Val Ser Met Glu Thr Gln Glu Ser Ser Ser Thr Leu Thr
260 265 270
Val Ala Ser Arg Asn Ser Pro Ala Ser Pro Leu Lys Glu Leu Ser Gly
275 280 285
Lys Leu Met Lys Ser Ile Glu Gln Asp Ala Val Asn Thr Phe Thr Lys
290 295 300
Tyr Ile Ser Pro Asp Ala Ala Lys Pro Ile Pro Ile Thr Glu Ala Met
305 310 315 320
Arg Asn Asp Ile Ile Ala Arg Ile Cys Gly Glu Asp Gly Gln Val Asp
325 330 335
Pro Asn Cys Phe Val Leu Ala Gln Ser Ile Val Phe Ser Ala Met Glu
340 345 350
Gln Glu His Phe Ser Glu Phe Leu Arg Ser His His Phe Cys Lys Tyr
355 360 365
Gln Ile Glu Val Leu Thr Ser Gly Thr Val Tyr Leu Ala Asp Ile Leu
370 375 380
Phe Cys Glu Ser Ala Leu Phe Tyr Phe Ser Glu Tyr Met Glu Lys Glu
385 390 395 400
Asp Ala Val Asn Ile Leu Gln Phe Trp Leu Ala Ala Asp Asn Phe Gln
405 410 415
Ser Gln Leu Ala Ala Lys Lys Gly Gln Tyr Asp Gly Gln Glu Ala Gln
420 425 430

14/122

ccg act gct cct ggc tct gtt ggc cct cct gat gag tct cac cca ggg Pro Thr Ala Pro Gly Ser Val Gly Pro Pro Asp Glu Ser His Pro Gly 525 530 535	1754
agt tct gac agc tct gcg tct cag tcc agt gtg aaa aaa gcc agt att Ser Ser Asp Ser Ser Ala Ser Gln Ser Ser Val Lys Lys Ala Ser Ile 540 545 550 555	1802
aaa ata ctg aaa aat ttt gat gaa gcg ata att gtg gat gcg gca agt Lys Ile Leu Lys Asn Phe Asp Glu Ala Ile Ile Val Asp Ala Ala Ser 560 565 570	1850
ctg gat cca gaa tct tta tat caa cgg aca tat gcc ggg aag atg aca Leu Asp Pro Glu Ser Leu Tyr Gln Arg Thr Tyr Ala Gly Lys Met Thr 575 580 585	1898
ttt gga aga gtg agt gac ttg ggg caa ttc atc cgg gaa tct gag cct Phe Gly Arg Val Ser Asp Leu Gly Gln Phe Ile Arg Glu Ser Glu Pro 590 595 600	1946
gaa cct gat gta agg aaa tca aaa gga tcc atg ttc tca caa gct atg Glu Pro Asp Val Arg Lys Ser Lys Gly Ser Met Phe Ser Gln Ala Met 605 610 615	1994
aag aaa tgg gtg caa gga aat act gat gag gcc cag gaa gag cta gct Lys Lys Trp Val Gln Gly Asn Thr Asp Glu Ala Gln Glu Glu Leu Ala 620 625 630 635	2042
tgg aag att gct aaa atg ata gtc agt gac gtt atg cag cag gct cag Trp Lys Ile Ala Lys Met Ile Val Ser Asp Val Met Gln Gln Ala Gln 640 645 650	2090
tat gat caa ccg tta gag aaa tct aca aag tta tga ctcaaaaactt Tyr Asp Gln Pro Leu Glu Lys Ser Thr Lys Leu * 655 660	2136
gagataaagg aaatctgctt gtgaaaaata agagaacttt tttcccttgg ttggattctt caacacagcc aatgaaaaca gcactatatt tctgatctgt cactgttggtt tccagggaga gaatggggag acaatcctag gacttccacc ctaatgcagt tacctgttagg gcataattgg atggcacatg atgtttcaca cagtgaggag tctttaaagg ttaccaa	2196 2256 2316 2363
<210> 34 <211> 662 <212> PRT <213> Homo Sapien	
<400> 34 Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser Pro Arg Thr Leu Arg 1 5 10 15 Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg Arg Lys Val Lys Gly 20 25 30 Lys Glu Gln Glu Lys Thr Ser Asp Val Lys Ser Ile Lys Ala Ser Ile 35 40 45 Ser Val His Ser Pro Gln Lys Ser Thr Lys Asn His Ala Leu Leu Glu 50 55 60 Ala Ala Gly Pro Ser His Val Ala Ile Asn Ala Ile Ser Ala Asn Met 65 70 75 80 Asp Ser Phe Ser Ser Ser Arg Thr Ala Thr Leu Lys Lys Gln Pro Ser 85 90 95 His Met Glu Ala Ala His Phe Gly Asp Leu Gly Arg Ser Cys Leu Asp	

17/122

cctgtaatcc	cagcacttcg	ggaggccaag	gcagggcgat	cacgaggta	ggagatcgag	2340
accatccctgg	ctaaaacggt	gaaacccgt	ctactaaaaa	ttcaaaaaaa	aattagccgg	2400
gtgtgggtgg	ggcaccctgt	agtcccagct	actcgagagg	ctgaggcagg	agaatgggt	2460
gaacctggga	ggcagagctt	gaggtaaagcc	gagatcacgc	cactgcactc	cagccctggc	2520
gacagagcga	tactctgtct	aaaaaaaaaa	aaaaaaaaaa	aaagaaaata	agtgaacaat	2580
aaatgtaatg	tggctgaatc	attccaaaac	aatccccca	ccccagttca	cggaaaaatt	2640
ctcccacaaa	accagtccct	ggtgccaaaa	aggttggga	ccgctaattct	aaataatcta	2700
atcttcattc	aatgtaaaaa	aatgaataaa	ctttttttta	aatacacgt	ctcactttgt	2760
tgcccaggct	ggagtacggt	ggcatgatca	cagtcactg	tagctcaat	caccaggcc	2820
ccagcgttcc	tcccaccaa	acttctgag	tagctggac	tacaggcagc	caccaccatg	2880
cccagctaat	ttttaaattt	tttatagaga	tgggggtctc	accatgttgc	ccagactgg	2940
ctcaaaccct	gggctcaagt	gatcctccct	caaactcctg	gactcaagt	atcctccctc	3000
cttggcctcc	caaagtgtctg	ggattacaag	catgagccac	tgtacccagc	tggataaaca	3060
tttaagtctg	cactacatgc	atggacaatc	aggtttaatc	acatgcagta	tggacagtga	3120
gtccccagggt	ctgctttcc	atactgaaat	acatgtgata	ctaaggagaa	aggtgtcgc	3180
aaggatattt	aaaatgaaga	atatttaaaa	tgaggaaaaaa	actgttttt	catgactttg	3240
ataaggctga	taaagaccat	ttctgtgatc	tcaggtgatt	cactcaagta	gtatatttca	3300
gtaatcattt	tctggAACAG	cctgaattt	aaccaaaaata	ccatgatttt	ttaatgctgt	3360
tatgatacct	tgatgatatg	acccaaatgc	aatgttaggc	gctaaatctc	cacgagttt	3420
acttccccga	gagttgacag	ttttcttcac	aaattaaaga	aatatatttt	ttgatacatg	3480
attggcatat	ttaaaaacta	cactgaaatg	ctgcaaaatg	atataaaagaa	acattttcca	3540
gaatcaaatg	caatcaaaga	gtggattagg	aatttactca	ccattatcaa	ctaaatagaa	3600
acacttggac	tgggtgtgg	ggctcacatc	tgtatctca	gcacttggg	aggccaaggc	3660
agggtggattt	cttgaggcc	ggagctcaag	accagcctga	gcaacatagc	aaaactctgt	3720
ctctacaaaaa	aaaatgggg	attaaccagg	catgttgc	gatgtttgt	atcccagcta	3780
ctctggaaagc	tgaagtagga	ggactgttg	agcccaggag	atcaagactg	cagtggccg	3840
tggtagtgc	gcccacagc	ctgagtgaca	gagagagacc	ctgtctcaa	aacaaaaaca	3900
aacaaaaaac	acttaacctt	cctgtttttt	gctgtttgt	tttgttgg	ttttttttga	3960
gatggagtct	cactctgttg	cccaggctgg	agtgcgtgg	cgtgatctt	gtcactgca	4020
agctctgcct	ccccgggtca	cgccatctc	ctgcttcagc	ctcccagta	gtctggacta	4080
taggcggcc	ccaccacgccc	cggtactttt	tttgattttt	tagtagagat	ggggtttac	4140
cgtgttagcc	aggatggtct	tgatctctg	acctcgtgt	ccacctgcct	cggcctccca	4200
aagtgttgg	attacaggca	tgagccaccc	cacccggcca	acctttctgt	tttttagttt	4260
gatatgtttt	ttaacttcagc	agctgaaaga	atgctgaaag	tggccttcag	taaaaaatt	4320
tcactagaat	cttacatcc	atatttaatc	tgaatgcata	tccagatgt	tcaatgtt	4380
caaaaacact	catcatcatt	cctgtatgacc	tctaattctg	gtttcggtt	tctatttcaa	4440
tggaaacaga	ataaggaaag	aatggaaagg	gctctggaaa	tttgtccgg	gtatagata	4500
ctatcaaaga	tcaccaacaa	taagatctct	cctataaaata	taaaacaatg	ataattaatt	4560
ttttttaattt	tttttctct	tcagggattt	tttttcaag	ataaaacata	acttctaccc	4620
atacttattt	ttccaaagggt	tagaaaaatg	gttttctc	atctttatct	tcaagaggt	4680
cacagcaatg	caaacatcta	taaaatgcct	ctgcataatt	gtcagaagct	atagtcaga	4740
aatcattgaa	aatgcctttt	cattttaaagc	ttaggtgagg	tgttttagga	acctctatg	4800
acaacttact	ctattttattt	ggaggttaaac	tcccagactc	tcccagggtc	tcctgtattt	4860
atcttcattt	tttaggcttcc	taatcccttg	aagcacaatc	aaaaaaagcc	tggatcttt	4920
tttcgcacat	atcatecgcc	aatttcattcg	gttccagca	agctgacact	ccatgataca	4980
agcggcctcg	cccttcctccg	gacgcccagtc	tttgcgtccgg	ttagcttaga	tgaggggttt	5040
gctgggttc	agtgcaggct	tctgcgggtt	cccaagccgc	accaggtggc	ctcacaggct	5100
ggatgtcacc	attgcacact	gagctcttgg	caggctgtac	caatttttta	attattttat	5160
attttttttt	aaaatttatgg	tgaatattt	gttattctgc	tctaaaatag	gccctaaaat	5220
gcacagcaga	tatcttttgg	aacccacaggc	tttccacttgg	aagaactaag	tatttttctt	5280
ttaaaagatgc	tactaagtct	ctgaaaatgc	cagatctct	accttttcc	atcccaaact	5340
aagacttgg	atttatgaga	gatctagct	acagaaaatcc	cagacacatc	attggttctt	5400
cccagagtgc	agtccctccta	aagaggctca	gcctctaagca	ggcccttc	ccaggagggt	5460
gggtctgaga	cccacatagc	acttcccaag	gtgcgtctc	cagagaggca	ctgaaaacagc	5520
tgagcacaag	cttgcaagcc	tggagaactc	tcacagttag	aacggagggg	gcccaagtgg	5580
actaacatata	agagaaaaagg	gaacacagag	aaatggatgg	caccaacaa	cagcaagcc	5640
ttcatggcca	atgaaagcat	cagtgcacgg	gccagaaccc	tcatccccaa	agactcttca	5700
ctgcctttag	tgaaaaacaa	tggctagaga	gtgaagttat	gatcatgtat	agagaggtaa	5760
agttacatctt	ttatattctg	actctgttac	tgtgaaattc	cctatctgt	agactaaaag	5820
tttcagacac	cctgttcaaa	tatcccattt	gttgctagag	acttaaaatg	aacagaacgc	5880

20/122

tactaaaaat	acaaaaaaatt	agccgggcgt	ggtggcaggt	gcctgttagtc	ccagctactt	13140
gggaagtctga	ggcaggagaa	tggcgtgaac	ccaggaagcg	gagcttgcag	tgagccgaga	13200
tcacgcact	gcactccagc	ctgggtgaca	gagaagact	ctgtctcaaa	aaaaaaaaaa	13260
aagactcatt	tctctagatc	ttgagccgta	ttcaaattta	tctcagctt	gtgagaggtt	13320
aaagcaagga	atatccttcc	ctgtgggccc	tgctccttac	tgaaggaaagg	taacggatga	13380
gtcaaggaca	ccaatggaga	aaagcactaa	caccattatc	tgatgaacat	tacgtgaaga	13440
agggttaagaa	gtgaagtggaa	attgctgtgaag	aagtcaagtga	aagcggagat	tcatttgggg	13500
aaatggataa	tagggaaatcc	ataaaatgt	ttaaaaatgg	tttagaggt	gaggcgggggg	13560
gaccacaggg	tcaggagatcc	gagaccatcc	tggtaacac	ggtgaaaccc	catctctact	13620
aaaaatacaa	aaaatttagcc	aggcgtggt	gcagggcacct	tgatccccaa	ctactcgggaa	13680
gactgaggca	ggagaatggc	atgaacctgg	gagacggagc	ttgcagttag	ccgagatcac	13740
gccactgcac	tccagcctgg	gtgacagagt	gagactccat	ctcaaaaaaa	aaagtttagat	13800
acgagagata	aagatccacaa	agacacacaa	ctgctaattc	tgaacagaac	aaaacaatg	13860
gcacaggaaa	agaaaattta	agatataaca	ccggaaaact	ttccctgaat	tgagtaactg	13920
aatctatagc	ttgaaagggt	ttagcatatg	ccaaagaaaa	tcagtagagt	ccaaccagca	13980
caagacacat	ctagcaaggc	tggtgattct	accaacacag	agaaaagaatg	gggtgaccca	14040
taatgcggaa	aaaggcagac	catctgcagt	cttctccaga	acactgggt	ctgaagacaa	14100
aagaatgctg	cctactgagc	cagaaggggag	agaaaatgtac	ccaacacatc	tttaccaagt	14160
tagaatgtca	cgcattattt	aaaggctgca	aaagccatga	aagacatgaa	agaaacacaag	14220
catttacaaac	atgaaaagaac	acaagcattc	tcatactcaa	gaatccttaa	aaaaaatgt	14280
gtcctaattcc	agccccactga	aagttaaatg	tacccatgt	gctcattaat	gggaacttca	14340
tagcttcaaa	tcagttcttgt	cccatctacc	aacatctctc	gccccggctt	cctgcaatag	14400
tcagcacctt	tcctctctcc	cagtttgc	ccctggagtc	tgctctcagc	atagcagagt	14460
gaccacatca	acacccaaatg	cagagccctc	cagtgcgcac	tggtctacaa	agcccttccc	14520
acccccccacc	ccacgtgccc	tccggatct	tgtgacgtgt	tcctctgcata	ccctagcagc	14580
cctggctcc	tcactgcccc	tcctgtacat	caggaaggcg	actccttgc	tcttggctct	14640
ggccgcctcc	tccacctgca	gtgagttAAC	tcccttacct	actcttagtgc	attgctcaaa	14700
tgtcagccatc	tcacatggggc	cctccctgac	tacccttattt	aaattctaca	tactccccc	14760
gaccccatgg	acccctactca	cccttatttca	cttttatttct	tacaatttag	cacttgttct	14820
cttcttaacgt	attctaaagac	ttactcattt	attacattgt	ttgccacccc	ctctagtagca	14880
taaactccat	agggggcaggg	atttctgtct	atttattcat	ttctttatcc	ctaggacata	14940
gaacagggca	tagttcagag	tattcaatgt	tatcaatgaa	tgaacttagca	gtagtagccag	15000
ttcccaatgt	gcacacaaat	aaatctaaat	agaattaaat	ctcatgttct	gggttaacta	15060
tggatagaaa	atttagatata	attttagaa	gccttagaaag	aaaaaatata	taatgtaaaa	15120
ataatattaa	tttgataata	ataacaaaaaa	ctctgcccagg	cactgttgc	aaatctgca	15180
atcccaccta	ctcaggaggg	tgaggtggaa	ggatcacttgc	agaccagat	tcaagactca	15240
gcctaggcaa	cacggcaaga	aactgtctct	aaaaaaatata	aaactttaat	ttttaaaaaa	15300
gaattctcaa	agcgtcacaa	aaactggaga	ttaaggtaca	ggaatgtga	agtaatatta	15360
ctatgtcaat	ggtttttttt	tttttttagaa	aggataacc	aaaagatttc	tttctcaagt	15420
cgataaaactg	agaaaagataa	gcatatcttc	caattaaacag	agggggggga	aaaggccagat	15480
acaacaaaat	aagatataaaa	ttagtttcca	tttgaaaaca	agagtaggg	ttattttgca	15540
tcacccctacc	tgtgacctcc	cccagcccaa	aaaacactac	tgataaaacag	ggtagaaaag	15600
catcatctca	gataaaggcag	aaaaaactgc	cacagtctca	aaccacaaac	tataagcaca	15660
cacctggcca	accctgcca	gtctgggtc	agtagggaga	acgtgtctgc	agcttaggt	15720
taccaactt	gacatttctgt	gggatacaga	tgttccctgg	gggttcacac	catctcaaaag	15780
gcacctgtaa	tgccttactga	ttacagccac	catatgttag	agagaaactc	aggccactta	15840
gagagtataa	caagaacactt	atgtcatctg	agatgaggaa	tcctcagccc	tgcaattaa	15900
ccaaactctt	agaacaactg	gaaaaacata	aatatccaca	actttttgtt	cagtaattcc	15960
actcttagat	atcaatccaa	agtacatgt	acagcagata	cacacacaaa	atggtattta	16020
ctgcagcatt	gtttataata	gaaaaaaaca	agaaataatc	catatgttc	aataggatac	16080
tgggtacatg	agggtatgt	cccatcattc	aaccatcaaa	aaaggtgata	tggatgtcca	16140
cagatggaca	taaaaaagctg	tgtgttacgt	gaaaacaaac	tcaagcagca	gcaggatggg	16200
ctttagatag	tcagttatgt	ctaatttctg	aaaaaaaat	tcttagtgcgt	gcacagaaaa	16260
catctgaaag	aacagaaaaca	aaactatctag	cagaatattg	agatgtttt	ctaagttgt	16320
tatctatact	gtttgttata	tttaccccaa	gcaagaatata	ctttttggaa	aaaaaaaatt	16380
caggaatataa	agcattttctt	taaaatttcat	tttttttttt	tttttttttt	tttttttttt	16440
ttcttattca	tcataaaacac	acacagcaca	catgcacgc	tgtgcgttag	cacacccttt	16500
acttgataaa	taccatgttg	aatatttttag	tctttctttt	taggttctat	cccttcactc	16560
aaaatgcggt	tataaataaaa	tgtacttttc	atgtgccttc	tgcctaaacc	cactttaata	16620
taactttaca	gtccccattat	cattatagtc	tcaaaagctg	actcagcctg	aaactaccct	16680

22/122

acccaatatg	aagaacagaa	agaaaaataga	ctggccaaa	aataaagaag	aaaaaagagg	20340
agcagcagga	ggaatgtgg	aaaaagagaa	aggaagggaaag	gaagggaaagg	agggaggggaa	20400
ggagtgaggg	agaaaagtctc	aaagacctct	gagactaaaa	taaaagatct	aacacttgc	20460
atcagggtcc	aggaaagaga	caaagatggc	acagctggaa	acgtattcaa	aaaataatag	20520
ctgaaaactt	ccccaaatttg	gcaagagaca	taaacctata	gattcgaaat	gctgaacccc	20580
aaataaaaaag	cccaataaaa	tcccacccaa	aatacatcat	agtcaaactt	ctgaaaagac	20640
gaaaagagaa	aacgtcttga	aacgagttag	tgaacaaca	cttcatgtat	aaggaaaaaa	20700
caattcaagt	aacagattt	ttacagaaat	taaggaagcc	agaagggaaat	gacacaatgg	20760
tttcaagtg	ctgaaaagaa	agaagtgtca	acacaaaatt	ctagattcag	taaaaatatc	20820
cttcaagaat	caatggggaa	tcaagacagt	ctcagataaa	gaaaataaag	agaatatgtt	20880
gccagcagat	ctccccctaaa	ggaatggcaa	aagaagatc	atgcaacaga	ccaaaaaatg	20940
atgaaaagaag	gaatccagaa	acatcaagaa	gaaaagaata	acatgtatgg	aaaaaataca	21000
tgttaatttca	ataaaatttc	tatctctct	taagacttct	aaatttatat	gtgggtgaa	21060
gaaaaaaaaat	taaccctgtc	tgaagtgtt	ctactaaatg	tatgcagaga	attataatg	21120
ggggaaatgt	aggttctat	acctcattga	agtggtaaaa	tgacaacact	gtgaaaagtt	21180
acatacacac	acacacgtaa	gtatataaa	atatatgtgt	gtatatgtgt	gtgtatata	21240
atataatatac	ataatgttaat	acagcaacca	ctaacaacac	tatacaaaaga	gataataacc	21300
aaaaaaacatt	tagataattt	gaaatggat	tctaaaaaaat	attcaataac	tctacaggaa	21360
gacaagacaa	aaagagaaaa	aaagaggagg	acaaactaaa	ttttttaaaa	acataatataa	21420
aatggtagac	ttaagcccta	acttataat	aattacataa	atgttaatga	tctaattata	21480
tcaattaaaa	gacagagata	gcagagttaa	tttaaaaaca	tagctataag	aaacctgttt	21540
tgggctgagt	gcagtgactc	acacttgtaa	tcccgacact	tcgggaggcc	aaggcgggtg	21600
gatcaccgtg	ggtcaggagt	tccagaccag	cctggacaaac	atggtaatac	cccatctcta	21660
ctaaaaatac	aaaaaaattt	gccaggcatg	gtggcacacg	cctgtatgc	caactactca	21720
ggagggtgcg	acacaagaac	tgcttgaacc	cggcagcag	aggttagcagt	gggccaagat	21780
tgcgccactc	cagcctgaac	gacagagtga	gactccacct	cagttaaaa	acaaaaaaaga	21840
aacctgctt	aaatatacca	acatatgtt	gttggaaatta	aaagaataaa	atataatcatg	21900
aaaacattaa	tcaaaaagaaa	ggagtgcata	tattaataac	ataaaataga	cttcagagaa	21960
aagaaaattt	caagagacag	gaataaaaagg	atcaaaaaaa	gatcctgaaa	aaaaaagcagg	22020
caaataatc	attctgttg	gagattcaac	accctcttct	aacaactgat	agaacaacta	22080
gaaaaaaaaa	tcaagcatgga	gttgagaaga	acttaacacc	actgaacaaac	aggatctaat	22140
agacatttac	ggaacactct	acccaacaat	agccaaataa	acattcttt	caagtattca	22200
ctgaacatcat	ccttagaccc	taccctgggc	cataaaaacaa	agctcactag	tgattgcccga	22260
aggcttggat	ggacagtggaa	agacgtcat	ggggagggag	aagggtacag	ttaaagagt	22320
taggatttct	ttttgggata	atggaaatgt	tccaaaatttgc	attgtgttga	tgttggcga	22380
actctacaaa	tataaaaaaaag	gccatttgaat	tgtacgtttt	aagtgggtga	aacatatggt	22440
atgtggatta	tatctaacgc	ttttttaaaa	cttaacacat	ttcaaaagaat	agaagtctata	22500
cagagtgtgc	tctacttgaa	tcaaaactaga	aagaggttaac	tggaggataa	cgagaaaaagc	22560
ctccaaataat	ttgaaaacttgc	gacagcacat	tctaaaatc	atccgtgggt	caaagatatt	22620
cattttctgtat	attcatatttt	attgtttat	gtatttttaa	aaattttctta	agggaaataaa	22680
actgactaaa	aatgaatatg	gctgggtgcg	gtggctcacg	cctgtatcc	cagcactttg	22740
ggaggccgag	gctgggtggat	cacaagatca	ggagttcgag	accagccctgg	ccaaagatgg	22800
gaaaccccggt	ctcaactaaa	aaactacaaa	aagttagccaa	gcgcagttgc	gggagccgt	22860
ggtcccagct	acttggggagg	ctgaggtagg	agaatcgctt	gaacacaggc	agcagaggtt	22920
gcagtgcggc	aaagattgtgc	cactgcacgc	cggcctgggc	gacagagact	gcctaaaaaa	22980
aaaaaaaaaa	aaaaagaataa	tcaaaatttg	ttggacatag	ttaaagcaat	gctgagaggg	23040
aaatttataa	cactaaatgt	ttacattaga	aaagagaaaaa	agtttcaat	caatagtctc	23100
caactccatc	tcaagaacac	agaagatgaa	gagccaaaata	aacccaatgc	aagccaaaaga	23160
aagaaaatat	aaaaaaatataat	cgtaaaaatt	gaaaacagaaa	acacaataaa	gaaaatcagt	23220
gaaaacaaatgt	actgtatttt	cgaaaagatta	ataaaaatttga	caaaccctta	gcaaggctaa	23280
caaacaaaaaa	agaaaagaaga	cacggattac	cagttattag	aatgaaagca	taatttagaaa	23340
caactctaca	cattataaaat	ttgacaatgt	agatgaaatg	gactaattac	tgaaaaaaca	23400
caaattacca	caactcaccc	aatatgaaat	agataatttg	gatgcctga	taactactga	23460
gaaaatttggaa	ttttaattt	taacacttctt	aaaacagaaa	cattaaactt	aatattttat	23520
aaatattttaga	taaggttaattt	atacccttcc	ttaaacaataa	aaaacgacaa	attattttgc	23580
agctaaagag	atgtatgtac	tgtggaaaat	atcttcagaa	aaatagaact	ttgtttgaag	23640
aataaggatt	taaaaaatgt	ttttaactct	caagaagcaa	atatctgggc	ccagatggtt	23700
tcactgttgc	attcttacca	atgtttatg	aagaattacc	accaactcta	catagcatct	23760
ttgagaaaaac	tgaagagaag	ggaacatctc	ccagttcatt	ttatgaagtg	ggtgttactc	23820
tgataactaga	actgtataag	gacagctact	cttgacacac	tgcctatggg	tagtctgtct	23880

24/122

actgtcattt	gaggctgaca	ataaggcacac	ccaaagctgt	acctccttga	ggagcaacat	27540
aagggttta	accctgttag	ggtgttaatg	gtttggatat	ggtttgttgc	gccccaccga	27600
gtctcatgtt	gaaatttgtt	ccccagtact	ggaggtgggg	ccttatttgg	aggtgtctga	27660
gtcatgggg	tggcatatcc	ctcctgaatg	gtttggtgcc	attcttgcag	aatgtgatgt	27720
gttcttactc	ttagttccca	caacaactgg	ttataaaaaa	cagcctggca	ctttccccca	27780
tctctcgctt	cctctctcac	catgtatct	caactggttcc	ccttccttgc	atgcaatgag	27840
tggaaggcgc	ctgaagccct	cgccagaagc	agatagtgt	gccatgcttc	ttgtacagcc	27900
tacaaaacca	ttagcccaat	aaacctttt	tctttataaa	ttatccagcc	tcaggtattc	27960
cttatagac	agacaaatga	accaggacag	ggggaaatca	acttcattaa	aataatctat	28020
gcagtacta	aacaaataag	aacaagaggc	tccagaatgt	ggaaggccat	acccagagtt	28080
cctacaatac	atgtatctaa	aagtccagtt	tccaacccaa	aaatatataat	atacaggccg	28140
gacatggtag	cttatgtctg	taatcccagc	actttggat	gctgaggccg	gcagatcacc	28200
ctaggtcagg	agttcgagac	cagcctggcc	aatatggcaa	aaccggctt	ctactaaaaaa	28260
tacaaaattt	agccaggcat	ggtggtgat	gcctgtatc	ccagctactc	gggaggctga	28320
ggcagggaat	cacttgaacc	caggaggcag	agttgcagt	gacccggat	cacccactg	28380
aactccagcc	tggcaacaa	agtggagactc	cacccaaaaa	aaaaaaaaaa	tatacatata	28440
tatatgtgt	tgtgtgtgt	tgcgcgctg	tgttatata	cacatacaca	tatacatata	28500
tatacagaca	cacatatata	tatgaagcat	gaaaagaaac	aaggaaggt	gaaccatact	28560
ttctgtgtt	atgtatggat	gggttacac	ggggaaatga	gacaagggaa	actgcaagt	28620
agagacaaac	gtttagatgt	ttaacagaaaa	aaacttttg	agtaaccatt	ataaatatgt	28680
ccacagaatt	aaagaaaaagc	gtgattaaaa	aggaaagga	aaatattata	acaatattac	28740
tccaaataga	gaatataat	aaaggcatag	aaattataaa	atataatata	atggaaattc	28800
cggagttgaa	aggttagata	actaaaattt	aaaatttact	agagaagggt	caacactata	28860
tttgaactgg	cagaagaaaa	atttagtgag	acaatataat	ttcaatagac	attattcaaa	28920
tgaaaaataa	aaagaaaaat	gaatgaagaa	aaataaacat	aatctcagca	aaatgtggca	28980
caccataat	cacattaaca	tatgcatact	gagagtaccg	gaagcagatg	agaaagagga	29040
agaaaaataa	ttcaaatgtat	ggccagtaac	ttcttagatt	tttggttttaa	agcaataacc	29100
tatacaatca	agaaaactcaa	tgaattccaa	gtaggataaa	tacaaaaaaag	accacaaaca	29160
gatacaccat	ggtaaaaaatg	ctgttaatgc	aaaacacaga	aaatatttggaa	agcagctaga	29220
ggaaaactta	taagagaacc	tcacttacaa	aaagacatca	cttataaaaa	aaccacaaata	29280
atagaaaaac	ttgaccttc	atcagaacata	atgaatgtata	acatattttga	agtgtctaaa	29340
gaaaaaaaaat	aaagattctc	atatacgcaca	aagctgtctt	tcaaaaaat	acatccaaaa	29400
ggattgaaac	cagggtcttgc	aagagttatt	tgtacatcca	ttgttcatagc	agcattattc	29460
acaatagcca	aaaggttagaa	gcaacccaaag	ggtccatcgaa	caaataataa	aaatgtggta	29520
tatgtatata	caatggaaat	tattcgtat	aaaaaggaa	tgaatttctg	acacatgtca	29580
caacatggct	aaacatgttag	aacactatgc	taagtggaaat	aaggccagcc	aaaaaggaca	29640
aataccata	tacttcaact	gtatgaaata	cctagggttag	tcaaaatttgc	agatagaaag	29700
taaaacagtg	gttgccaaagg	gctgagggag	ggagtaacgt	ggagttatttgc	ttgaatgggt	29760
acagaatttgc	agttttgcac	gataaaaaaa	gttctggaga	cagatggtgg	tgagggtgg	29820
acaacaataat	aaataatactt	tataactatgc	aaacttataac	ttaaaaatgt	ttaacatgg	29880
gaaaccccg	ctctactaaa	aataaaaaaa	aattagctgg	gtgtggggc	gggcacctgt	29940
aatcccagct	acttggggagg	ctgaggcagc	agaattgttt	gaaaccagaa	ggcggagggtt	30000
gcagtggatc	gagattgcgc	caccgcactc	tagcctgggc	aataagagca	aaactcccg	30060
tcaaaaaataa	aaaaataaaaaa	aaaatttataa	aatgttataa	caggaggcc	ggcacgggtgg	30120
ctcacaccta	taatggccagc	actttgggg	gcccaggcc	gcatcactt	gagaccaggaa	30180
gtttgagacc	aggcctggcc	acatggccaa	accctgtctc	tgctaaaaat	acaaaaattt	30240
gccaggcatg	gtggcatata	cttataatcc	cagctactgg	tgagactgag	acacgagaat	30300
tgcttgaacc	caggaggcag	agattgcagt	gagtcgagat	cgcccaactg	aattccagcc	30360
tgggcgacag	agcaagattc	tgtctcgaaa	aaacaaaaac	aaaaacaaaa	agcaaaaacca	30420
aaaaataattt	aagcaggaaa	cgagattgt	gctgaggagg	agaaagatgt	cgaggacc	30480
ggctcatgg	agcacaatgg	ttttcaaaaa	atgtttaatg	attaaaatgg	taaattttat	30540
atgtatctt	ccaaaaaaa	aagggtgggg	ggcaggaaaa	tgaagggtgaa	ataaaagacat	30600
cccagagaaa	caaaagttaga	gaattttgtt	ccttagaaga	aacaccacag	gaagttcttc	30660
aggctgaaaa	caagtgaccc	cagagggtaa	tctgtatct	cacagaaaaat	tgaagcatag	30720
cagtaaagg	tattctgtaa	ctatgcact	aaacatgtat	atttttctt	ttcttctctg	30780
aaatgattt	aaaagcaattt	gcataaaaaa	ttatataata	agcctatgtt	tgaacctata	30840
acatataat	aaatataat	gtatataatt	tgcaaataac	tgccaaaaag	agagttggaa	30900
caaagctttt	actaggctaa	agaaattact	acagatagta	aagtaatata	acagggact	30960
taaaaataaa	attttaaaaa	atttaaaaat	aataattaca	acaataat	ggttgggtt	31020
gtaatattaa	tagacataat	acaaaaataac	cacaaaaagg	gaagaagaca	atagaactac	31080

ctctctca	gacaccctc	caaggctgc	gagcacgtgc	catgctatgg	ctttctccaa	41940
catcaggaaa	tgttctccac	tcagttcac	cttaatacaa	atgtgttctc	tcttcagaga	42000
aggcaaaaaa	attcatgacc	atctgactgg	gagaagtcat	ttcttagtaa	agtgtccatc	42060
tttttctgag	gaacacagga	ggaaaatctt	acagaaaaga	gttaacacag	caggcctaag	42120
actgtttttt	aaaataaaata	aataaataaa	taaataaaata	aataaataaa	taaataaaata	42180
aataaaatgaa	tgatagggtc	ttctgtattt	gccaggctag	tctcaattt	ctggcttcaa	42240
gagatcctcc	cacccatggc	tcccacagt	ttgggatttt	agacatgagc	catttgtctt	42300
ggcccaagac	tgttttctt	aaaagtctc	ataaaaagca	tggtaatcc	ttggctggca	42360
cctggaaact	tagatttcg	aagggttccc	accatccaac	ctggaaagag	ggactcaactg	42420
tgcctaaattt	atttgtgtgg	ttatgctgaa	cttctgctt	tcttcaggta	gcgtggaatg	42480
tggtatgtgc	ttggcaaaagg	gggcctgcat	gaccagcccc	caataaaaac	cctgggggtt	42540
gggtctctag	ttagtttttt	ttgttagacag	catttcacat	gcgttgcac	agctccttcc	42600
tcggggagtt	aagcacatac	atccctgtgt	actgcactgg	gagaggatgc	tttgaagctt	42660
gtgcctggct	ttctttggac	ttggccccc	gacacccatc	ctttgtgtat	tgtgttttgt	42720
atccttcac	tgtataaaat	tacagccgt	agtagacaccac	atgctgagtc	ttccaagtga	42780
accaccagat	ctgagcatgg	tcctggggc	ccccaaacaca	gaaataaatt	ataaaagacc	42840
aaggactggg	catgggtggcc	catggcggt	atctcagcgc	tttgggaggc	cgaggcagga	42900
ggaccagttt	agccccaaaag	ttcaaaatgtt	cagtgcacca	tgactgcgc	aatgcactct	42960
aacctgggg	acagagcaag	accctgtccc	caaaacaata	aactaaacac	ataattctgc	43020
cttccaagt	tcttaaaatt	caatgaaatg	gtagaaacat	ttttaaaaca	ctaaatcaaa	43080
agaaaacctgg	aaaacaagag	tgccgatggc	caactaaaat	gtctagaaa	tttctgaaaaa	43140
gtaaaaatgtt	ctcagaaacca	gattactgtt	gcaaaaccata	gccaataaca	agcttgggag	43200
gaggctgtt	tgcagaagga	aatggtaaca	gttttccagg	aacagacttg	taacagcaga	43260
tagaacagca	gaggtagaaac	ctgacaagg	gattacctgg	ggaactgcag	tctgaaatgac	43320
caggactgtt	ggacccttcc	cctcacaatgg	aatacacacg	ccactcagca	gcacaccaca	43380
gctcttcaac	aatcacagga	ggcacgctac	gcttagtaag	acaggaaaaa	aggaattctc	43440
aaacttcgaa	gatgaacaca	taaagaaatc	ccaaatgtttt	attcgtatg	atgaaacacagg	43500
gacactgaat	caacagaaca	caaaccaaag	caaagataat	tactagacca	catagaagaa	43560
attatttagat	attcttggga	agacctaagg	ggacattata	aagagcaagc	atgtggatatg	43620
tgacgatctt	tgtgtatatac	caagaaataa	aaacacagga	tgaagaccag	atagagaata	43680
atgctactat	ttgtgcaaaa	aaggagaaat	ggagaatctg	attcatattt	gcttgttattt	43740
gcatgaagaa	acttttggaa	gtacataatg	aactaacaac	aatggttacc	tacttgtaa	43800
gcgagagaag	taagaggaca	ggaatggtt	gaacacccctt	tgtgtccgga	atttgggtt	43860
tcttggctgt	acttggagaa	tgaaggctgt	gaccctcgcc	gtgagctaa	caggctttaa	43920
aggcggtgt	tctggagttt	gttccttctg	atgtttggat	gtgttccgag	tttcttcctt	43980
ctgggtgggtt	cgtactctcg	ctgactcagg	agtgaagctg	cagacccctcg	cggcgaggt	44040
tacagcttctt	aaagggggcgc	atctagatgtt	gttcgttcc	cctgttgcgt	tcgtgttctc	44100
gctagcttca	ggagtgttgc	tgcagaccc	cgaggctgtgt	gttgcagctc	atatacgac	44160
tgcagacccca	aagagtgttgc	agataataaga	acgcattcc	aacatcaaaa	ggacaaacact	44220
tcagcagcgc	ggaatgcgcac	cgcagcacgt	taccacttcc	ggctcgggca	gcctgttttt	44280
attcttttat	ctggccacac	ccatatctgt	ctgattggtc	cattttacag	agagccgact	44340
gctccatttt	acagagaacc	gattggtcca	tttttcagag	agctgtttgg	tccattttga	44400
cagatgtctg	atttggctgt	ttacaatccc	ttagctagac	acagggtgt	gacttgggtt	44460
tttacaatcc	tttagcttaga	cataaaatgtt	cctaagttcc	caccagactc	aggagccctag	44520
ctggcttcac	ccagtggatc	cggtcatgt	gccacagg	gagctgcctg	ccagtcggc	44580
gccctggcc	cgcaactcc	agccctctgg	ttgtcgatgg	gactggcgc	cgtggagcag	44640
gggggtgggt	tgtcaggggag	gtcgggccc	cacaggagcc	caggaggtgg	gggtggctca	44700
ggcatggccg	cccgccagg	ataggcgctc	ccccgcagg	aggcgactaa	ggcccagcga	44760
gaaatcgcc	acagcagctg	ctggccagg	tgtcaagccc	ctcactgcct	ggggccgttg	44820
ggggccgt	cccgccggcgt	cccaatgtt	ggcccgccaa	gcccacccc	acccggaaact	44880
cacgcgtggcc	cgcaaggcacc	gcttacagcc	ccgggtcccc	cccgccctc	tccctccaca	44940
cctccctgca	aagctgaggg	agctggctcc	aggcttggcc	agccccagaaa	ggggctccca	45000
cagtgcagcg	gtgggctgt	gggtcttcc	agcgccggca	gagtggcgc	taaggctgag	45060
gaggcaccga	gagcgagcga	ggactgccc	cacgtgtca	cctctcaattt	tatatttatgc	45120
ctttttataa	cagtctgtt	ttgaacactg	attatcttac	ctatttttt	tttttttttt	45180
tgagatggag	tgcgtctctgt	tcgcccagac	ttggatgtcag	ttggtgccttc	ctggctca	45240
gcaagctcc	cctcccccgggt	tcacaccatt	cttctgcctc	aacctcttca	gtagctggga	45300
ctacaggcaa	tcgcccaccac	gccccagctaa	ttttttat	tatttttttt	ttagtagaa	45360
cggagtttca	ccatgttagc	cagatgttgc	caatcttcc	acctcgatgat	ccatccgcct	45420
cgccctccca	aagtgttggg	attacagacg	tgagccacty	cgccctgcct	atcttaccta	45480

29/122

tttcaaaaagt taaactttaa gaagtagaaa cccgtggcca ggcgtggcgg ctcacgcctg 45540
taacccccgc actttggag gccgaggcgg gccgatcactg aggtcaggag atcgagatca 45600
tcctggtaa cacagtaaa ccccgctcg actaaaaata caaaaatta gccgggcgtg 45660
gtggggca ccggcagtc tcgctactgg ggaggctgag gcaggagaat ggcgtgaacc 45720
tgggaggcag agcttgcagt gagccgagat agtgcattg cttccagcc tgggcacag 45780
agcgagactc cacctcaaa aaaaaaaaaaa aaaatagaga cccggaaagt taaaatatg 45840
ataatcaa tttaaaaaca ctcagagat ggctaaaaga gttgacggaa caaatctaaa 45900
ttagatgg tgacatcgca aaaccgc gaaacatc ccagaatcgca gcccataaag 45960
ataaaagagag catttcgc gggcacagtg gtatggcagg ggaattgcct gagtccaa 46020
gttcaggc acattgaacc acaccattgc actccaggcc tgggcaaac acgaaatactc 46080
tgtctcaaa aaaaaaaaaa tttaattaaa aaagacagaa tatttgagag aaaaaaaaaatgc 46140
ttatttcaag aaacatgaaa gataaatca gatattctaa ttcccaagta agaataattc 46200
cagaagcaga aaatagaata gggcaagga aacactcaaa acttctccag tgccatagaa 46260
atgtgttta atctttagaa tgaaacggac tccaaatgc tgagcaggaa gaacaaaaga 46320
gatccactct taagccagtg tggtgcggca ggcgtggc tcatgcctgt aatcccagca 46380
cttgggagg ccgaggcagg tggatcacct gaggtcagga gtttggatc agttagggcca 46440
acatggtcaa accctgtctg tactaaaaat acaaactttaa gctgggtatg gtgggcaca 46500
tctgtatcc caactacttg ggaggctaa gcaaggaaat cacttggaaac caggagggtgg 46560
agggtttagt gggccgagat catggcacac tcccgccctg ggtgacagag caagattcca 46620
tctcaaaaaa aaaaatccact cctagacaaa taatagttaa attttagaac accaaggaga 46680
aagaaaaaaa attgtaaagc ttccagagaaa ataaacatttta actacaaaaga aacgagatc 46740
agacgcgtgc acttcttcct agataccgc agataaaagca atatctccaa aattcagaag 46800
gttttaacgt agaatccctat accccatcaa gaatattcac atggaaaatg gaaataaaaaa 46860
acattgttta aacatgcaag gtttcagaaa gtttaccat cacagatcc ctggaaaacaa 46920
aaccaataaa tcacttaagg actcattaaag aaaacaaatg aaataaaaagc accaatgtatg 46980
agtaaataat cagaaaaatt tacagttac ctaaataact gtttatgcatt aatgtatgaa 47040
aacccaaaaa ttaatatgg gacagaatta aatcatgtat aagattttt tttgttttac 47100
tcatggagag ttccatataaag catattatc tttaatagca agaaaaaaa atgttttagat 47160
atgtgtgaaa aactaagggt accaaaacag tgcaatttca ttatcatca ggaaaatcca 47220
aattaaaaacc acagtatcca ccagaataac taaaaggtaa aagacagaaa ttaccaagag 47280
ttggcaagaa tggggagcaa ccacatatac ttctgggta aataaggttg tgcaaccgg 47340
actgaaaaact gtttgcgtat atctactaaa accgagcaca tgccacagact acaaccaagc 47400
agttccactc ccagatcac actcaacaa aatgcacaca ctcactcaac aaaagacgtg 47460
tactagatg ttcatgtact tactattcat aatagtccaa aatgcacaa aaccaactgc 47520
caatcaaaatg ccaatgtata tctatattag ggtatatac aatggcatat acacagcaat 47580
gagaatgaaa tgaaccagct cggcacagtg gttcatgcct gtaatctcag cactttggc 47640
gggttaaggca ggcagatcac ttgaggctcg aatatttggaa ctggcttgc caacacgg 47700
aaaacctgtc cccactaaaaa acacaaaaat tagccgggca tagtgggtgc aggctgtaa 47760
ttccagctac tggggaggct gggttgggaa aatcggttgc acccgaaagc cgagggtcgc 47820
agtggcgaa gatcgccca ctgcactca gcttggacga tagagcaaga ctccgtctca 47880
aaaaaggaaa tcaaaaaataaa aaaaataatg gacaggaata atccgaaaaa gatcgtata 47940
caaaataat ataaatgggc taaagctacc tattaaaaga caaagatttc acacccataa 48000
ggatagctac tatccccaaa agagagagaa taacagatgt tagcaaggat gtatggaaac 48060
tgtatttc acgcattgtc ggtggaaata taaaatggt cgccttgc ggaaaaacact 48120
atgctgggtc ataaaaataaa taaaatagaa agtactactt gatccaaacaa ttctacttct 48180
gggtatatac ccaaaaact gaaaggcaggc tcttgaagag atatttgcatt acccatgatc 48240
atggcagcat tattcataat agctatgtat tgaaccaac ataaatatcc ttgtataat 48300
atatggataa gcaaaaatgt gtgtatatac tcaatggat attatattgc aataaaaaatg 48360
aagaaaaattc tgacacatgc tacaacatgg atgaaccttgc agggcattac attaaatgaa 48420
ataagccagt taaaataaaagaa caaataatcat atgggtact atattagata ctcatgc 48480
gtacctaaaa taggcaaaattt catagagaca aaaagcagaa tgggtggtgc cagggctgc 48540
ggtatggat acagagcttc aattttgtaa gatggaaaaaa ttctggagat tgggtgcata 48600
acaatgtgca cacacttaac actggggacat tggaaacttca aaagtagtaa atggaaaaaa 48660
taaaaaataat aaataataaaat ttttatgttca ttatccaca atatttattaa aagacaaaag 48720
attaactaat taaaataaaat ccagccatataa gctaatggta aggtacacaa taaaagaaga 48780
cacagaaaaat tggaaaatcg tggactagaaa aagatattcc atataatgc taacaaaag 48840
caagtacagc aatataaaga gaatgaacaa aaaaaaaaaat aaataagatg gctgtttat 48900
tcccaaaaagg tacaattcaca caagaagata caagaatttgt gaacctttaa gcacataaaa 48960
cagcttcaaa aatacaacat taaaagaaaa atatattata aacatagaaa tagtacaaaa 49020
acccttacaa gaatcataat gggagtcttc aatacaactc tccatataa caggtaaac 49080

30/122

agagaaaaaaaaa	aataagttaa	ggatgcagaa	aacctgaatt	accatcaata	aacttgagat	49140
taatatagaa	ctgtatacc	aatatactaa	gagttcaggg	aacagtctg	actgacagt	49200
gactgcaat	taatctttc	ttaatctttt	tttttctttc	agcaactgtt	cagaatagag	49260
atcctaaaaaa	ccttccagct	acaaaacatc	ttttaaaaaa	tataaaaaaa	tacaaaaata	49320
actctgaaat	caatagaaga	cacatggtga	aacccaaaatt	ctagaataca	gggagaataa	49380
aggcatttc	agatattaca	aaaacagaaa	attgtatcatt	gctgaagtaa	tttctaaaga	49440
atgtacttga	gggagaagaa	aaatgttcca	aagaaaagta	tctgtatcac	aagaaggaaat	49500
ggaaagtcaa	gaaatggtta	acaggtagat	aaagctaata	aatgttgacc	tagaaaataa	49560
caaaaacaat	acgataatag	tctcggtga	agggttgaag	taaaaataca	attaaggcca	49620
aatgtgaggt	aagtggat	aaagaattag	aagtccctgc	cttgttcaca	ggactgatta	49680
aataaatgag	ccaggtttc	cattcaaaaca	gttaaaactt	gaacaaaata	aactcaaatt	49740
aagttagaaag	ataaaaaaaca	gaaattatg	tcatagaaaa	ataaaaaatc	aatagaatta	49800
atcaataatc	cctggttat	aaaagctgtt	tcttggaaag	gattaataaa	ataatcatta	49860
agcaagtctg	atcaaaaaaa	aagaaaaaaag	gtacaaaaaa	aagtactgt	tcagaaagag	49920
aacatacaga	tacatacaga	tatgttaagag	tctgttttct	tacaccagaa	tactatatac	49980
aacattatgc	tagcatatat	taaatttcaa	taatgttaat	gattttcttag	gaaaacagaa	50040
aatattaaat	ttactttgaa	gaaacagaaa	aactgagaaa	aataatagat	catgaaaaaa	50100
atgaaaaggat	aattaaatcc	tgatattac	tgcctttaaca	acaccagcag	cagcccagc	50160
agtctgcagt	caagttctgc	caaaacttgag	ggaacagata	attcttctat	tccagagcat	50220
agaaaaatgt	gaaaaggttt	ccaaatttaat	cagagaggac	agcctgatcc	ttgttatgaa	50280
cacagataaa	atggggtaa	actatatgcc	aaactcagat	acccaaaccc	taaataagat	50340
gctagcttat	tgatgtgaa	aatccaaaag	tgcattttaa	attagcccg	ggttttagag	50400
aaagaaaatc	tagcaatgt	accaccactt	atgttaacaa	tttaagacg	aaaatctaca	50460
tgatcatatc	aatgcatgt	acaaaaaaagc	atttgggca	aaaacccaaac	accaccctt	50520
gacttttaa	actcttagta	attaggcata	aacagaaaatg	tacttaatgt	gatagaatac	50580
actcgtgaa	gatacagagg	gaatgctccc	taaaaccaag	cccaagacaa	agattcctat	50640
ttaacctcaa	tagtcaacac	tgcagcgaga	gtaatctatg	gaagacaagg	aaaaaagtaa	50700
aaacatgaga	gacatctgtt	tttaacacaga	caataagatc	acctaacttgg	aagaggcaaa	50760
cgaatcaagc	ggaaaaactat	taaaacttgag	acaggctta	tgatggaggc	tcagcttcag	50820
ctgttagttt	ggcttacaaa	ttcaacttcgc	ttgcttggag	agttatcc	gcaaagctaa	50880
tttctgttga	ggtatttagga	ttgacaagcc	tgtgctctc	cctccccc	catcttcaac	50940
actgaaaataa	cacgggtttt	ggaacttgat	aacagaatct	tccaaaacaa	aaaattgtcc	51000
tgaagggtctg	acttgttcccc	ttactcaaaa	aacactttat	ctgtgcctg	cagctccctac	51060
agttgctgtt	ggataaggct	gccaacccgc	tccgcgtat	tcttcctgca	gaggggcaagg	51120
aagagcaactt	tcacagggaa	attttttcc	gaactgtatg	ccgatttata	cataaaactta	51180
cgtgtggca	aatggagctc	cagcaaaaata	agatattcag	agtcaactt	ccttaggaaa	51240
aaaaaaaaaaa	aaaagcaagc	acataaacat	aattttcttg	catgggcaact	ggggaaaggag	51300
gtcggttactt	ccgcacgc	gcaggccgc	accaccggga	aaccacccgg	caccgcgc	51360
tgccccccgg	ccttccaggt	gcactgcgc	gcccgcggcc	agctgaccgg	ggatgcgcag	51420
ccctagccct	tccccctgtca	ccccggccag	gaagggggcg	gagcgcggcg	gacgcggagg	51480
gcgaagggtt	tctcggtctt	ctgcaccacg	cagcacccccc	aaggcacaac	aggagggtt	51540
cgggaggctc	ccgagaccca	ggagccgggg	ccgggcgtgc	ccgcgcacct	gtcccactgc	51600
ggcgagggtt	ggggtcgccct	ccagggccgc	agctgtcg	agccacccgg	ctctcagttcc	51660
cggttccctg	cgacacccct	cgggccggga	ggggaggagg	cgccacccgt	ccgtgtccac	51720
ctggggcacc	ggtccccacc	ctccggggcc	ggcaggacag	gccaggacgt	ccctcctggg	51780
ctggggacag	gacacgcgc	gaggggacc	gggccccccgc	ggcgaagacg	cagcacgc	51840
tcccaaaaaa	gcaatcccgt	gcccccacga	cgactgc	gaccccccgc	ctcgcccccc	51900
catcccttca	gaccacgc	ctgaggcg	aaagagccggc	cgccgggggg	gtggccggcg	51960
cggttagtac	tcacccggcc	cgctgtca	gcccggccgc	aaccccccacg	ggcacggct	52020
cgccggcgtc	actgtatgc	aggagggaa	cccgccgtcc	ggccggccgt	ccagccatcg	52080
ccgcacgggg	gcgagcgc	gcccgcgg	gtcgctgg	agatgtatg	cccgacccgc	52140
ccgcctgc	gtccctcttc	agccggccgc	cgggggcccc	ctctctccca	gtctcagtg	52200
tctcatctcc	ctatctgtc	atccctgtgt	cgcacataat	cgatgtttgg	gcgtcccaag	52260
ccagatgtgg	accccat	cgcaacttac	actggaggtt	ttctaaagggt	ggtccccgg	52320
ccagcagctt	cagcttac	tggaaatgt	agaaaaatgc	gattctccgt	cccaacccagc	52380
ctatccgtt	tttctgtac	taaaaccatg	aaagggtgggg	ccagcagtcc	acattctgc	52440
aagcccg	agtgtatct	aggccccc	cagtttggaa	gtatgtca	cggcctcacc	52500
tccggccccc	aaggagcccg	gtcttgc	tggcgctagc	cgcacacgg	cacccatcc	52560
tgcggggccc	gccccccccc	tgcaccc	ccgccccaaacg	cctccctccgg	gatgcagegg	52620
aggcgccctgg	aagtccggca	ggtcaacatc	ccccctcagca	tcttccctac	cctcacggct	52680

31/122

cctccctccag	gggtgcctca	tggccagggg	ttagaaagag	ccactgtgtt	tcttgacatg	52740
gaagtggcct	aagaccttaa	tgaaaactgc	aggagtggaa	tgacagaacc	tttggtcata	52800
cttgaggggcg	tgaagctcaa	atgaggagga	agggaaaggat	ccagggagaa	taaccaaccc	52860
tggcaagtgc	tggcgccccag	gtagaggggc	gagcctaggc	tagcggttct	cgaccaggc	52920
cggtgttgc	cctcctcgcc	gccccgcgta	catttggga	ggtctggaga	catttttgtt	52980
tgtcatgatg	cgggagttgc	tactgttgc	taagtggta	gacacgaggg	tgctcctcaa	53040
catcctacct	gaaggacagg	actgcccac	aaccaatgg	gatccggccc	caaataagaa	53100
accctgggtc	ggtcagcaac	aacccctttg	ttctgagaag	agagggagaa	agaataaaaag	53160
aagtgggtg	aagtttttgtt	ttggtagagg	aaacttgcag	acattttccac	tggaaaggaa	53220
gagaggaga	ggaggggagat	gtctgttaagg	acgagcaaac	cgggtgacag	ctgatttcct	53280
catattgaag	taatgagtcc	tagttataat	aaattcctaa	aaaaaaccca	gttattccct	53340
gcaataaaact	tgtctttttt	ttttaataat	actgtctgtat	tctgtttgt	aatattttat	53400
ttacaggctt	tgcattgtata	tgcaaaaaatg	agatggggca	taattttctt	tttgaatgtc	53460
taatgttgtt	tggtttgcaga	atcaatgtta	tgctcacatc	ataaaaaattt	tggaaaccgag	53520
gcaggaggag	tgcttgaggc	cagaagttcg	agaccagtct	aggaaacaca	gtgagacccc	53580
cccatctcta	aaaaaaaaaa	aaaagaaaaa	aaaatggca	tgtttgttctt	ttccttttac	53640
tctgaacaat	ttaaggagca	ttaaaaattt	ctattctttg	aggtttgtatc	attttccagt	53700
aaaaaaaaatgtt	cctcccgagcc	tgatgtttc	tttggggagg	gtaaatcttt	taaggctaga	53760
aaagttttctt	ctgtggcaat	tttattttt	acatttaaa	aattttctta	gaggtaattt	53820
tgataaaagca	tgtattttctt	aaaacaaaattt	atcctttttt	tccagatgtt	caagtgtatt	53880
tgcataaaagt	tgagggaaagt	agtctttgt	gaatctttt	acttctccca	aatatcttat	53940
tttgtgtatt	tttgcttctt	tatttgtta	actttttaaa	gtgtattttt	tttcaaaaga	54000
atcagctctt	aggttttatgt	ttttgtttat	actggagctt	tttttttttt	tttttttttt	54060
tatttttctt	cctttttttt	tttagacgtat	tttgcattaa	cgtaatcgga	agaaggtaaa	54120
tttagaatctt	ttgttactat	ttgtttttt	tttctccctt	tttctctgaa	gtctgtttt	54180
ataaaatagta	ccatgttatt	tgtgcataaa	tattcatttg	tcttatattc	ttggaaattt	54240
tcccacttca	tcataaaaatg	accttccttg	tctcattttaa	tgtgttcaaa	cttgcctgt	54300
aatttaactt	tgtctgtat	tttacccatc	tgctgaattt	tggtttttac	cccaaacaac	54360
cttgcgttt	ttcgtctttt	ctgaaccctt	tattttaggt	aatcccttga	attagagcac	54420
taagtttgc	tttgcatttt	aatctgaaaa	tctttatctt	gccatagatg	agttgagccc	54480
tattcatgtg	acagctatat	tatgtgttt	catagccctt	ttggtccctt	tttactcttt	54540
gcattgcata	ttttgtgttt	attgtgtttt	gtgtttctt	tgataatttgc	gaagggtttgt	54600
atttttatcc	agggagttgc	cttataatca	tacttccggca	tacacatgt	cctcagtttgc	54660
ttcagactgt	ctgtttaactc	cctattctga	ataaaaatgt	cattgttatt	ttccctttttt	54720
ttcttacccc	ctttttctt	cctcacctaa	tgtaaatgtat	tttatcttc	tttagtatttt	54780
gctttttaa	ttaactacat	ttataaaat	cttatactact	tgattttttaa	atcagcttttgc	54840
aatgagatat	ttggattcct	agatataaaa	gatgttaattt	ataccatcc	cacgttagta	54900
ggtttataaa	atcatacatc	ctgctgttgc	accataatcc	cacgttttttgc	ttatgtttccac	54960
tcctacagtt	aaaagattca	gaagtattat	taacagttat	tttgccatag	ttttttcccc	55020
aaccctttt	gtggtaagg	atgatccgc	tttagtttct	taagaataat	ttatagagca	55080
gaggtgtgt	gctcacgtt	gtaatcccag	cactttggga	gacaagaggt	agaaggatcg	55140
cttgaagcca	gcagttcaag	accacccctga	gcaacatgt	gagacccctgt	ctctacaaaa	55200
aattttaaaa	tttagccaga	cgtatggcg	tgtgcctata	gtcccagcta	ctcaggaggc	55260
tgaggcaaga	ggattgtctg	agccccaaag	tttgaggctg	cagtgcaccc	tgattgtgccc	55320
actgcacccc	agtcggcga	agaaatgtag	aacctatctc	tttttttttt	caataataaac	55380
ttatgaaaat	tatattccct	gagttttca	tttttttttt	tttttttttt	tttttttttt	55440
taaaagtttgc	agtataaaattt	tttgcatttt	actttttttt	tttgcatttt	tttgcatttt	55500
gtctcttgc	attgggttttgc	gtctgtat	gaaatggcgat	cagctgttttgc	tattttttttt	55560
cagaaatgg	gtatgtccgc	gcccggacc	gtggctcatg	cctgtatattc	caacacttttgc	55620
ggaggccgag	acagggttgc	cacgaggctc	ggagatttgc	accatcttgc	ctaacatgtt	55680
gaaaccccccgg	cttactaaa	agtacaaaaaa	gttagctgg	catgggtgt	gacgcctgt	55740
atcccagcta	cccgccggagc	tgaggcagg	gaatggcgat	aacctgggg	gaggagcttgc	55800
cagagagctg	agatcgccgc	actgcacttc	agccctgggc	acagatgt	actccgtctc	55860
aaaaaaaaacaa	aaaaaaaaacaa	aagaatgtaa	gtatgtccca	tgatgtccca	agaattatct	55920
ttttgtcttat	gaaatccaga	aatctactgt	tttatacattt	tttgcattttt	tttttttttt	55980
aatattttctt	ggggacacaaat	agatttgactc	tatagattta	tttttttttt	tttttttttt	56040
acagagtctc	actgcatttct	cagcttactg	caacctctgc	ctcacgggtt	caagcaattt	56100
tcctgcctca	gcctcccaag	tagctggac	tacaggcgc	tggcaccatg	cctggctaat	56160
ttttgtctttt	tttagtagaga	cagggtttca	ccatgttgc	cagggttgc	ttgaacgcct	56220
aacctcaagt	gatccaccc	cctcagccctc	ccaaagtgtct	gggattacag	gggtgagcca	56280

32/122

ccatgcccag	cctcaattcc	tctttctatc	tggttaattt	tctgaagttg	aaaacattt	56340
ttctaatacg	ttatttcagt	gttcttctaa	gatgtgtaaa	gcaccctatt	cccaggtcag	56400
ccccatctt	gctagtgagc	tcggctgggt	cttcacaaga	gctctgggtt	tctcctgttt	56460
aatctcaagt	acctctgtca	gcctccacct	ggtttatgt	ttggaggttt	ttggttttt	56520
tttttgttt	ttgacagagt	cttactctgt	cacccaggct	ggagagcagt	ggcataatct	56580
cagctcactg	caacctctgt	ctcccagggt	tgagcgattc	tcctgcctca	gcctactgag	56640
tagctggat	tacaggcgcg	tgccaccaca	cccggtcaat	ttttgtattt	tttagtagaga	56700
tggggttca	ccatgttgc	caggggtggc	tgttaactcct	gacctcagggt	aatccacctg	56760
cctcagcctc	ccaaagggtct	gagattacag	gctgtgagcca	ccgcgcctgg	catggtttgg	56820
agtttaatc	tgttagttta	ataaaagatag	tgcttatgtt	tgtgtttctt	atatttctt	56880
gtactcttgg	gtaatttgtt	agatccccat	atctacacaa	gaagtccatt	ttcaatttctt	56940
ttcttcagac	tgtttatttt	attttatttt	attttatttt	tatgttttag	atggagtctc	57000
gctgtgtcac	ttctggagggc	tgtagtgca	tgccgcgatc	tcaggtcact	gcaacccctcg	57060
tctcccggtt	tcaagcaatt	ctcctgcctc	agcctcccg	gtagctggga	ttacaggcac	57120
ctgcacattt	ttaatttttt	tagagacaga	gtctcgctt	gttgaccagg	ctggagtgcg	57180
gtggtgcaat	catggctgac	tataacctcc	aaatcctggg	ctcaagtgtat	cctcctgcct	57240
cagccctctg	agttagctggg	actacaggca	catgccacca	tgcccaggtt	attttaattt	57300
ttttttagag	acagggtctc	catatgtgc	ccaggctggc	ctcctactcc	tggcctcaag	57360
taatcttcct	acctcagcgt	cccaaatttac	taggttata	agcatgagcc	accatgcccc	57420
gccttgcct	actactttaa	tttcatatgt	taggtgacca	tgtaattttagt	catccaaacc	57480
aggatactgt	aagaatgaaa	gaggctgaca	gtagttatgt	gctgggacta	gcattgtgca	57540
ctgagattat	ttctgggaaa	gcaggagata	cggtcaccct	acttataatgt	tgctttgtctt	57600
tggtgttgc	attttggagt	ttcttatttgc	aggcttattt	caactgggc	gccttgatcc	57660
gccctgccc	gcaatgtcac	cgttcttc	accgggtctc	tgggaccctt	tcagtcaacta	57720
tacttagctc	agttcccccac	cctcccaactc	cctaaaagcg	taaccaggaa	tcctgcctca	57780
ggtctactgc	cgttcccg	gggctgttcc	agtttctatt	acccagagtc	aaactcccg	57840
cattccctac	ctgattccag	acttggagtc	cagagcttta	acctcttcag	gccaactcccc	57900
cacttgcac	ttctgtccct	atatctttagt	ccatggagat	acatttcatg	tctttgagtc	57960
tacttacaaa	gtaaattttt	ctgtttttta	attttttttt	tgagatggag	tcttgcctg	58020
tcacccaggc	tggtgtgc	tgacccatc	tcggctca	gcaacccctg	cctcctgggt	58080
tcaagcgatt	catctgcctc	agcctccaa	gtagctgtga	ttacagacag	gcaccaccac	58140
gcccgctaa	ttttttttat	cttttagtag	agacagggtt	tcaccatgtt	ggccaggctg	58200
gtcttgaatt	cctgacccctg	tgatctgccc	atctcgccct	cccaaagtgc	tgagatttaca	58260
ggcgtgagcc	actgtgtcccc	gccaattttt	ctttttttat	atttcattgc	tatatgttta	58320
gaggataagt	tttacagtgt	atatgcattc	ccaaatattt	gacaaaaaaa	atccaaaaaa	58380
aatttagaaaag	aaaatccaaa	aaatctaaa	aaataccaaa	aagcaacaat	ctcacagacc	58440
atactcaactg	acccccaata	aaataaaatt	agaaaattaac	cacaacttta	caaaaataaaag	58500
tactcaatgc	agagaggaaa	gaggaaataa	acatccaaat	tacaaagtct	aggcgggtggc	58560
tcacgcctgt	atccccagca	ctttggggagg	ccaaaggccgg	cagatcacaa	ggtcaggaat	58620
tcgagaccag	cctggccaaat	atggtaaac	cccggttccca	ctaaaaatatac	aaaaatttagc	58680
caggcatagt	gtatgtgtcc	tgtatccag	ccacttggga	ggctgaggca	ggagaatcac	58740
tgaacccagg	gagacgaaga	ttgcagttag	ccaaaatcgt	gccactgcac	tctggcctgg	58800
gtgacaaaagc	gagactccat	ctcaaaaaaa	aaaaaaattac	aaactcttta	gatagaaatt	58860
ttggtgtttt	ttttggatggac	ggagtctcac	tctgtcgca	aggctggagt	gcagtgggac	58920
tatgtcagat	caccgcaccc	tccatctctt	ggattcaagc	aattctctgt	tctcagcctc	58980
ccaagtagct	aggattacag	gcgcacccaca	ccagaccagg	ctagttttta	tatttttagt	59040
agagatggtg	tttcacccatg	ttggccaggc	tggtctcaaa	ctcctgacct	caagtgtatcc	59100
acctgcctca	gcctcccaa	gtgtctagat	tacaggcgtg	agccaccgca	ccccacccctag	59160
atagaaattt	caacatgagg	ccgggcacaa	tgctcacgc	ctgtatctc	agcaacttcag	59220
gaggctgagg	cgtggggagg	tcacttgggc	ccaggagttc	aggacccaga	tggtgacag	59280
agacagaccc	tgtctctatt	tatttggaaa	aaaaaaaaaa	aaagagagag	agaaaagaaat	59340
ttcaacatga	aaagtatctc	tcaaaccctt	cgagatgttg	gcaaaaagcg	actccaaaggaa	59400
aaatgttata	ctgtgtgtga	atttgcgttga	aaataagaaa	gaggccgggt	gtggtggcta	59460
acacctgtaa	tcccaacact	ctgggagttcc	gaatcaagt	gatcatgagg	tcaggagatc	59520
gagaccatcc	tggctaaat	ggtgaaacccc	tgtctctact	aaaaatataca	aaaatttagt	59580
aggcgcgggtg	gctcatgcct	gtaatccca	cactttggga	ggctgaggca	ggggatcac	59640
ctgaggtcag	gggtttgaga	ccagccctggc	ctacatgggt	aaacctcgtc	tcttctacaa	59700
ataaaaaat	tagctggcg	tggtggtggg	tgcctgtat	cccagctact	cagaggctga	59760
ggcaggagaa	tcgcttgaac	ccgggaggcg	gagggttgcgg	tgagccgaga	tcgcaccact	59820
acactccage	ctgggcaaca	gcctgggtga	cacagtgaga	ctccatctca	aaaaataca	59880

aaaat tagct	gggtgtggtg	gcctgcgcct	gtagtccag	ctacccggga	ggctgaggca	59940
ggagaat gga	gtgaacctgg	gaggaggagc	ttgcagttag	ccgagatccc	accactgcac	60000
tccagcctgg	gcgacagagc	aagacttttgc	tctaaaaaa	aagaaaaaaa	aaggaaaaaa	60060
gaaccctgat	aataaagaaa	chaaragttc	aactctcaaa	gctcgacac	tttaaagaaa	60120
taattaataa	aggcagaagt	taaagggagg	atgataaaagc	aattttttt	gttgggtttt	60180
ttgagatgga	gtcttgcet	gtcaccagg	ctggagtgc	gtgatgcgt	cttggctcac	60240
tgcaacctct	gcctccggg	ttcaagcaat	tctccctgc	cagcctctg	agtagcttgt	60300
actacagggt	cgcgcaccc	ggcccgacta	atttttgtat	tttattaga	gacggggttt	60360
caccatattt	gttaggctgg	tctcaactc	ctgatctca	gtaatctgcc	cacctcgcc	60420
tctcaaagt	ctgggattac	aggcaggcgc	caccgcgcct	ggcctaagc	aaaatattgg	60480
ttctgtgcaa	aaggtaa	aaaagagcaa	acgtttacaa	actggagcca	gcacccattc	60540
agctcagtgt	gtctggagaa	aaaacaatc	cgcttcagaa	ttcatgatta	cgagccccctt	60600
tttgcctcct	aaaatccta	ctatgtgc	gtgaccatt	ctctctctt	ctctctctct	60660
tgcttcctct	ccagaaaagc	tattcagaca	ttctccctct	tcctcaaaacc	tccaacactt	60720
cctccatccat	ctttagcctc	agctgctgac	ctcacttcta	atcattgaga	aaccaggaga	60780
agcatttaag	agtgaaccc	cgccctcccg	cacgggcaaa	accacccacc	cacagaattt	60840
tgccccaat	ctgcgtccct	tccttcac	atggatggac	gttccaggct	ccgagccaaa	60900
gcccggctc	ccttggagct	ctggatccac	cacccgc	ttctcaggca	gggccccccagc	60960
agctccctgt	ctcccttgc	ccatcaatcc	ctcccttcac	tgggtcaactc	ccaacaatat	61020
atataattttag	tgatgttttct	cccatgttgt	aaaatca	agccttcctc	ctccccccagc	61080
tactatccta	tttgccttct	tccatttc	gaaaaacttc	tcaaagcatt	gtgtctatgt	61140
gtgactccta	tttatcttct	cccgcttc	gtgagtcct	ttccacagac	tctcacc	61200
gttactccat	gaaaatgcac	ctgactgc	acatccaaat	gtaatgttc	agttcttaat	61260
tttattcagt	ctttcagcag	catttgcct	ggccgatcac	ttcccttctt	taaaaataact	61320
tttctcagcc	aggcgtgatg	gctcacac	gtaatccaa	cactttggga	ggccaaggcg	61380
ggaggatcat	gagagccca	gagttcaaga	tcagcctggg	caacatggca	agaccctatc	61440
tctacaaaaa	ctaaaaagta	gccatgtga	ttgcatgcac	ctgtatccc	atctacttag	61500
gaggctgagg	cagttaggt	acttgcct	ggaaatcaa	ggctgcagt	agccatgatt	61560
gcaccactgc	actccagcc	gagtgcac	gagaccctgt	tcctaaagaa	caaaatagga	61620
aactttctc	agcatattcc	tctgatttc	ctgtgcgttc	tgtctgcaca	gattcagtct	61680
ccttgcggg	ttcttcctca	tcctctgtat	ctttgcac	tgaagtggccc	cagagtacag	61740
tctttttt	tttttttgc	acgcagtc	gtctgtacc	caagctggag	tgaatggcg	61800
aggtctcage	tcatgcacc	tctgcctc	gggttca	gattctctg	cctcagcctc	61860
ccaagtagcc	aggactacag	gcacatgcca	ccatgcccag	caaattttg	tatttttat	61920
agagacaggg	ttttactata	ttggccacgc	tggctcaaa	ctcctgaact	cgtgaaccac	61980
ccgcctcggc	ctcccaaagt	gctgagatta	caggcatgag	ccaccacacc	cggcccagag	62040
taagtcctt	agacggcc	tctactata	cttgcctccc	tcataaaactc	ctccctgcctc	62100
atggcttaa	ataccatcg	tagactgt	actccctat	ttctctttt	ttttggaga	62160
cggagtctcg	ctcagtc	caggctggag	tgcagtggcg	cgatctcggc	tcaactgcaag	62220
ctcacac	caagttcaca	ccatttc	acttcgcct	ctccagtagc	tggactaca	62280
ggcacccg	accacgcct	gctaattttt	ttgtat	agttagagat	gggttccacc	62340
atgttagc	ggatggtctc	gatgttctg	cctcgatc	cccccac	ggcctccaa	62400
agtgtggg	ttataggtgt	gagccacc	gcccagccg	tgactccat	atttctatct	62460
cggatgttct	ggggatgttct	ctcagaactc	catactata	aatccaa	tccataaaatag	62520
tatctcaa	ggcaatat	ctcaaaatgc	aatttctact	tttctcc	aacttgc	62580
cctgcagtct	ccaccat	aatgttcaat	cta	acattag	gaggcaaaaa	62640
attcttgc	tttcttct	acacacc	tccaa	tttcttct	ctgcagatcc	62700
caaattcc	atgttctc	catct	tttcc	tttcc	atcttctt	62760
ctcacctgaa	tca	tttcc	tttcc	tttcc	tttactcc	62820
tttagat	tttcc	tttcc	tttcc	tttcc	tttactcc	62880
ttccctgtct	gct	tttcc	tttcc	tttcc	tttactcc	62940
gagtcttcc	agt	tttcc	tttcc	tttcc	tttactcc	63000
tcactcc	ccag	tttcc	tttcc	tttcc	tttactcc	63060
tttagat	tttcc	tttcc	tttcc	tttcc	tttactcc	63120
ggctctct	tgca	tttcc	tttcc	tttcc	tttactcc	63180
ggccggacat	gg	tttcc	tttcc	tttcc	tttactcc	63240
tcac	tttcc	tttcc	tttcc	tttcc	tttactcc	63300
acaat	tttcc	tttcc	tttcc	tttcc	tttactcc	63360
gctgaggc	gagaat	tttcc	tttcc	tttcc	tttactcc	63420
cgccacaaca	cccc	tttcc	tttcc	tttcc	tttactcc	63480

aaaaaaatca	cacaaaacaca	cttctttca	tattctttt	ccaagttta	tttttctcca	63540
gaatacttta	cattgtttta	atggaaagtcc	tccgtttccc	cccaactaga	atggatactt	63600
cctgcaggta	ggcactctag	tcctccatc	caagtactaa	ccaggctcaa	ccctgcttag	63660
cttctgagag	cagggggagat	cagggctgtt	cagggtggtt	tggcccagga	attttgattc	63720
tgttttattc	attgctgttc	tgttGattct	cttttgttcc	tcctcctagt	gctgagaaca	63780
ctacttgtac	ataataagca	ttcaataaaat	atttgttcaa	tgaatgactt	gttgaatgaa	63840
ttaatcttcag	aatatcgagga	ctgggtctac	attagaaaaat	tttcaaggt	cattctctgt	63900
tgtcgtaaca	cattaagaga	ggaaaatttt	gtactctaa	tcatttgcata	aaatacatac	63960
tgatttctgt	tttcaaaaac	tcttagtggc	tggcgaggt	ggctcacatc	tataatccca	64020
gcattttggg	aggacgaggt	gggcggatca	cttgaggatca	ggagtttgag	accagcctgg	64080
ccatcatggt	gaaaccctat	ctctactgaa	aatagaaaaa	ttagccgggt	gtggtggcgc	64140
atgcctgtc	tcccgcact	ctggggaggt	gaggcaggag	aatggcttga	accggggagg	64200
cgagggtgc	agtggccaa	gatcatgcca	ttgacttca	gcctgggtaa	cagagtggaga	64260
ctccatctca	aaagaaaaact	cttagtgagt	tttaggaatcc	aaggaaagacc	ctcaaactaa	64320
atagataatc	tagctaccag	aaggcttcag	taaaccttaa	cactccatgg	tgaacattaa	64380
gaaacattcc	tactaaaaga	caggctaaga	atgcctgaa	tcttcacggc	tagtccaaga	64440
agtcaaaaag	aagaatggag	cgctgattta	aaaaaataaa	caaacaaaaa	actaccgatg	64500
cagagggctgg	cgaaggcagg	tgaaggactg	tacagtactt	gcctggagca	ggccggatggc	64560
cacaccctgc	cgaaggctgc	tcagctggct	gggggacgct	ccagtgtgt	agtggcagga	64620
tgcagggtac	ttcctctgc	aggaggatgc	actggggaga	tcctccccca	ctcacacttt	64680
ggcagctggg	gctttgaat	gtgacttagc	ttctgtcaaa	gggtcaatcc	accctttgat	64740
atatgatgca	aaggcgaaca	tatgtatc	aggtgagaga	acagccaaa	ttaggacttt	64800
taccacagct	gtggagggtgg	acagcgacag	tgttggggcc	tggccagact	tttcatgtctc	64860
aaagggtggt	gttggatcttc	ctacttttg	tccttcagg	gcttccttgc	cctgtgtgt	64920
gaacctgctt	cttttaattt	tttttaactt	ttttaaattt	ttaattgttt	taattaaaac	64980
aaattttgaa	aactgtctga	acctgtttt	gaaccctgct	atgatttggaa	tgttgtccc	65040
ctgcoaaact	gattttgaaa	cttaatttcc	aaagtggcaa	tatttgatg	gggttttaag	65100
cagtacttgg	atcatgagag	ctctgcactc	atagttggat	taatggatta	atagttgtc	65160
atggggatgg	catcgtggc	tttataaagag	gaagaattaa	gacctgagct	agcatgtgt	65220
ccccttcacc	atttgatatac	ttacactgcc	tagggctct	gcagagatc	cccaccaaca	65280
agaaggctct	caccagatac	agctctcaa	ccttgcactt	ctcagctct	gtaactgtaa	65340
gaaataaatg	ccttttcttt	atgaatttacc	cagtttcaga	tattctgtta	taaacaatag	65400
aaaacgaact	aaggccaaact	ctcatgatcc	tactgcctatc	ccatttccat	aaactccctt	65460
tatgtttaag	agagccagag	tggccaggc	gttggactc	acgcctgtaa	ttccagcact	65520
ttggggaggcc	gaggcagggt	gatcacaagg	tcaggagatc	gagaccatcc	tggtaacac	65580
ggtgaaaccc	cgtctctact	aaaaatacaa	aaaaattagc	tggcgtgg	agtgggtgcc	65640
tgtgttccca	gctactcggt	aggctgaagc	aggaggagaa	tggcgtggac	ccaggaggcg	65700
gagcttgcag	tgtgtcgaga	tcgtgcact	gcactccagg	ctgggtgaca	gaatggact	65760
ccgtctcaaa	aaaaaaagag	gccagattt	atttctgttgc	cttgcaacca	agaatatctgg	65820
ctgggtcact	gaagtttcc	taaataatag	caattttaaag	acttttcca	agccaggcaa	65880
tgccttgcct	tgtgtgttcc	ttgtggtaat	acatttattc	attcatttgc	tcaaccaact	65940
gtgcttccaga	gactaagaat	acaaaaatgg	ggccgggtg	tggtggctca	cacctataat	66000
cctagcaact	ttggggaggcc	aggcaggatc	atcaccttgc	gtcaggagtt	cgagaccaac	66060
ctggccaaaa	ttgtggaaacc	cctacttctac	taaaaatatac	aaaaattagc	tgggggtgt	66120
ggcggcagacc	tgtatccca	gctactcggt	agactgggc	aggagaatca	cttgaaccccg	66180
ggaggcagag	gttgcagtga	gccgagatcg	caccactgca	ctccagctg	ggcaacacaaga	66240
gcgaaactcc	acctcgaaaa	aaaaaaaaaa	aaaaaaagag	ggccggggct	gggcgcagtg	66300
gctcacgcct	gtaatccca	caactctgg	ggccaaggca	ggagaattac	gaggtcagca	66360
gatcgagacc	aggcctgacca	acatgtggaa	accccatctc	tactaaaaat	acaaaaattha	66420
tccggggctg	gtggcgcaca	cctctgtcc	cagctacttg	ggaggctgag	gcaggagaat	66480
cgcttgcacc	cgggaggcg	agggtgcagt	gagccgaaat	catgcactg	cactccagcc	66540
tgggtgacag	agtggacttc	cgtctcaaaa	aaaaaataaa	aaaaaaaaaa	gaattcaaaa	66600
attgttaggt	tatgtgtgc	ttcttagttt	gttggagagga	catctgttcc	tcaaggaaagg	66660
ctagaatcta	taccctgtat	ccttactgaa	atcaatccag	cagtcaaaac	atgggaccaa	66720
cgatcacagc	agtaagatag	gaagagcacc	tttgcgtacatt	tgatctatgt	tgagataagc	66780
cactgacaga	gctgaaggaa	gctcacagtt	ctgggttcca	tcctttggca	tttaaaaaaga	66840
aaagtgtctaa	aaaaatttgg	ttggtcacgg	tggctcacgc	ctgtatccc	aacactttga	66900
gaggccaagg	caggcagatc	acggggatcg	gagttcgaaa	ccagccctggc	caacatggtg	66960
aaaccccgtc	tctactaaaa	acagaaaaat	tagccgggca	tggtggcgca	tgcttataat	67020
cccagctact	caggaggctg	aggcaggaga	attgtttgaa	cccggggagg	ggaggttgca	67080

35/122

gcgagtgaga	gcaggccact	gcactccagc	ctgggagaca	gagcaagact	ctgtctcaa	67140
aaaaaaaaaaag	aaaaaaaaaaagaa	agaaaggaaa	aaaaggaaaaaa	aaaaaaaaaaagaa	aaaaggaaaaaa	67200
ttcagccag	gccaggccctg	gtggctaca	cctgtatcc	caacacttgc	ggaggctgaa	67260
gcgagacggt	gccttagccc	aggagttga	gaccagccctg	agcaacatag	cgagaccctg	67320
tctctataaa	aaaaaaatttt	tttttgcca	gacgcagtgg	ctcacccctg	taatcccagc	67380
actttggag	gccgagggcag	gtggatcagc	agtcaggag	atggagacca	tcctggctaa	67440
cacggtgaaa	ccccatctct	actaaaaaaat	acaaaaaaatt	aaccggcgt	ggggcggggc	67500
gcctgtatgc	ccagctactc	gggaggtga	ggcaggagaa	tggctgtaa	ccgggaggcg	67560
gagttgcag	tgagccgaga	ttgcggact	gcaactccaga	ctgggagaga	gtgagactcc	67620
gtctcaaaaa	aaaaaaaaaa	aaaaaaaaat	taattgtcag	gtgtgcggc	atgcagctgt	67680
agtcctagct	actcgggagg	ctgaggtaag	aagatcgctt	gagcccagga	gttcaaggct	67740
gcagtaatag	tgcctctc	tctaccctgg	gtgacaatga	gaccctctc	caaaaagaaaa	67800
aaaaaaaaaggg	aaagaagaaaa	agaaaggaaa	aaagagaaga	aaggaagggaa	gaaagaaaga	67860
aaaagaaaaa	aaaggaagggaa	agaagaaaaaa	aaaagaaaaaa	aagaaaagag	agagaagttc	67920
aaagacaaa	gggtcaggat	cccaaaatag	tttttatgtt	ttatttattt	atttacttat	67980
ttatTTTGA	gacagtatgg	ctctgtcgc	caggctggag	tgcagtgt	cgattgcggc	68040
tcactgcagc	ctccaaactg	ggctcagggt	gccctccac	ctcagcctcc	cgagtagctg	68100
ggaccacagg	cgcgtccac	catgcccgc	taattttta	attcttgc	gagatgagg	68160
ctctatatgc	tgcccaggct	ggtctcgagc	tcttggctt	aagccatcca	ccgcctggg	68220
cctcccaaag	tgcttggatt	acagaatgt	gccaccgcgc	ctaattgggt	ggtttgttt	68280
tttatttgcg	gggtctcgct	gtgcggcagg	ctggagtgcc	agtggctgtt	cacaggtgca	68340
gtcctggagc	attgcatacg	ctcttggct	ctagcgatcc	tccagagtag	ctgcagctgg	68400
gattccaggc	gcccacccgc	gccccggctca	gaatgggtt	ttatattgc	ggttatgtc	68460
ccacccatgg	gatataatgt	gtaccgaact	gtgtgcgcag	ggaggctgag	gttgcagtga	68520
gccaagatga	tgccaggcga	ctccagcgtg	gttgcacagag	caagattca	tctcaaaaaaa	68580
aaaaaaaaaa	aaaaaaaaaa	aagaattgaa	agtaaggct	tgaagagata	tttgtcctg	68640
tatggtcata	gcagtattaa	ctttgacc	ctagctaaaa	cacaaaagca	acatgtgtct	68700
gtcagcagg	gaacggataa	acaaaatgt	gtatataatgt	acaatttgc	attattcagc	68760
ctttaaaaag	gaataaaaagg	ctggatgcgg	gggctcagc	ctgtaatcc	aacactttgg	68820
gagactgg	ttgggtggatc	acccggagg	aggatgt	gaacagctg	gccaacatgg	68880
tgaaacttca	tctctactaa	aaataactaa	attagccggg	catggggca	cttgtctgt	68940
atccaaagcta	ctggggaggc	taaggcagga	gaattgttt	aactcaggag	ccggagggtt	69000
cagttagcta	agatggcacc	actgcactcc	acgctgggc	acagatggag	actccatctc	69060
aaaacaaaca	acaaaaaaat	tattatttc	aaagaaacaa	gaccctgggt	ccatttccca	69120
gcccacac	gtatgtgact	cacaacac	acgctgggtt	gctatgc	tgcttcattt	69180
aattgtcatt	ttaacttc	atcaccc	agtcctggaa	taactcttt	ctgaccttt	69240
tgtgtcgac	catctccatg	tcgctcaac	tgcagtcct	ctca	ctgcac	69300
gccagacgt	gtctgactgc	agggtcatcc	tttgtggctt	aggctgactc	gggcatagca	69360
gggtgtctg	agacccatcacc	gcatatagc	tttgc	ataaaactca	tataatattc	69420
atattatgt	gtcttggat	gtgtatctt	gactgtctt	ctcgatcag	tgccctcaac	69480
ctctttccca	ggatTTTCTC	ctctac	tc	ctgctctgca	aagacaaaaaa	69540
gctgcagagt	cccagctccc	tcctttacac	ccacgc	agcctctt	ctcagaacccc	69600
tttaaacaga	gtctttact	gcagatccc	aaacacggca	caccc	tcccacccac	69660
tccagacaca	cccgaggtaat	tatagcacc	aggtaacta	tgtatgtga	gtccctggaa	69720
catgtggata	gtggccccc	ggatgt	caaa	tgctggcacc	tgagagaac	69780
agggtgatc	ccaggaatca	gagcatggc	ctctgggagg	tagggatgt	gccaggcagg	69840
ctgccaaaaa	ttggtagagc	aaggccacag	gatcttctg	accttc	caaacagagg	69900
ctcctgtact	ggtgatccct	gtgttattt	accactcc	tcctgggggt	cgtgtctct	69960
gtccca	cccgaggactc	gtgtatgtc	ctactcgat	ccttttcat	agaagcatgc	70020
tgtcttc	cctgttggg	gcaagatgt	caatttca	actataatag	cagtggata	70080
cagagaagaa	aaaaatgt	gtggcaagaa	aaacaggctt	ccaagcagga	gttttctat	70140
aaaaacaaaa	acgttttacaa	gcaaaat	tataaagg	tagatgtaa	atattttagg	70200
ctttgagagc	cacatagact	tgtttgcagg	gactcaatgt	cgctatgt	gttggaaagc	70260
agccatcagg	gttatgtaaa	tgagtgtatc	tgatTTTGT	tcagcaaat	tttatttacc	70320
aaaacacagaca	atgatgtggc	tggatggc	ccatgtatc	tagtttgc	actcttgc	70380
tgggtcacc	cagatgtatc	tttgcattt	ggctctgt	ctggttagct	gcaggagctt	70440
ggaaggctt	ctgagcctgt	ttccttcatct	gtt	agcaataatt	tcta	70500
aagagtgtt	cctcacgcct	gtaatccc	cacttggag	gctgaggcag	gcccgtt	70560
tgagggtcaga	agttcaagac	cagcgtggcc	aacgtggca	aaccctgtct	ctactaaaaaa	70620
atacaaaaaa	tagccgggca	tggtgccg	catctgt	cccagctact	tgggaggctg	70680

aggcagggat	actgctagaa	cctggaggt	ggagcgtgca	gtgagtggag	atcacaccc	70740
cacactccag	cctggccgac	agagcggagac	tccatctcaa	aaaaaaaaaa	aaaaagagt	70800
ttagaagggtt	ttgagataat	gaataaaaaga	tgccttgggt	atactaagta	tccaacaact	70860
gatacgctca	ttggtctaatt	tataacagtt	tagaagcgat	tgagtcaaca	aatgctggat	70920
ttgtcaggga	ggacttccta	tcaggaggt	gatcttgggc	tgagtccctga	agcaaaagata	70980
ggcattggat	agaggagtt	agagaacacc	ctaggactgt	tattattatt	attcgacacg	71040
gagttcttgc	ctctgtcacc	caggctggag	tgcaagtggcg	cgatctccgc	tcactgcac	71100
ctctgcctcc	cagggtcaag	cgatttcct	gcctccctaa	tagctgagac	tacagggtgt	71160
tgccaccaca	ccccgctaat	tttttatattt	tttagtagaga	cagagttca	ccatgttggc	71220
catgtggtc	tcgaactcct	gacttcagg	gatccacc	cctcagccct	ccaaagtgt	71280
ggaataacag	atgtgagcca	ccgcacccag	cccagaacca	ttttcaatc	cttggctctg	71340
ccttttatta	gctgcaagat	ctcaggcaat	ttatttaaag	tctccaaaga	ctcattttct	71400
cattcacaca	atggggcaaa	taataatatc	tactatccca	ggtgttgc	agaattaaat	71460
gcaacatgac	atthaatgaa	atgagaagtc	ccttggacat	taactggct	aagtatgtgc	71520
tcgacaaggaa	tatcatttt	ggtggtact	tagcatctca	gaactgatgc	tcacaatgga	71580
atatcattga	aacgcattaa	aattcatttt	aatatgattgt	aggtatgtg	gcaattgaaa	71640
gaagaagaca	agaggactga	ttataatgt	tcaggctcac	tagtctctt	ttaggaggg	71700
aaaacaattt	caagtttaat	tttaggtct	agattttac	ccctgtct	cattagaatc	71760
accagatgg	atgaaatcag	agcccatctg	aggctgtgtt	tttcatctcc	agaatgagag	71820
ctgttgggg	gattaagtt	ttgaaaaagt	acatctaaca	ggtgatcgaa	aatgatagt	71880
atattattgc	agtgtatggc	attattgtt	ttattattat	actgaaagag	gttcagttt	71940
tctgatccat	aaagtgggg	aattgcatga	gaccattgt	aaatttctt	ctagctctgt	72000
tttttttttt	ttgtttttta	gacagagtct	ctgtcgcccc	ggctggagt	caatggcatg	72060
atctggctc	actgcaacct	ccgcctcccg	ggttcaaatg	atccctctgt	ctcagcctcc	72120
gaagtagctg	ggactacagg	cacacaccac	catgcccgac	taactttat	attttaata	72180
gagggtgggt	ttcaccat	ttgtcaggct	ggtctcaaac	tcctgaccc	aggtatgcca	72240
cccgctcgg	cctcccaaca	tgctgggatt	acaggcatga	gccactgtgc	ccaaacccctt	72300
ctagctttct	tgatcaactga	ttcttagggtt	ctctgtgaa	atattatgt	gacatcttgg	72360
ataaaagatc	atgcaagagc	tcccaatatg	gttataataa	ttgatcttgg	aggcttagct	72420
actcctgtat	gatttagacat	gactcaactg	cctctttat	gtgtacaaca	caacaacaca	72480
accaagaaag	gttattctgg	cattcattt	attcagttt	tttacagccc	ttacttccag	72540
cagcacgtt	aagatatggc	caggccccc	tgcaagtggct	caagtctgt	atcccaggac	72600
tttggggggc	caagggtggc	ggatcacaag	gtcaggaggt	tgagaatctg	gcaatttctt	72660
agacttagaa	gcaaccaggat	cgataacaca	gtcttgggt	ggctctccct	ctgtccctcc	72720
ctcgcttccc	tcatttctca	tcctcgcccc	ttgagactgt	caccttcaca	tagccctgcc	72780
atgagacctt	catctcaggc	tttgcattt	gggtaactg	aggctaaaca	ctgagtggcc	72840
ctaaaagagg	atgggattt	ggaagttaga	ttattccaca	gagaacagac	tttgtctatg	72900
atcggccca	ggttgttaat	gttggaaaaaa	agagaggatg	catagctta	tctcatctcc	72960
tagtcaaagt	caacaccatg	ataaaataga	gtcaaatcc	gagatgtgaa	ttggggacat	73020
ttgagtgggtt	aaccctgaga	agcttgcacc	ttcagacccc	tcaatacccc	tgctccccag	73080
agaaggctgg	acattgaccc	cagcacaggc	aggagccctg	caagatgcca	tttgtccctac	73140
taaaagatgga	ccccctccact	ctgtttctag	gtaaaataacc	aaagtcaagt	ctccacacag	73200
cctgagcaag	aaagtcaag	cctgtcatac	gagaaaaataac	cacactggcc	aaaggattca	73260
ctagccctgg	ccactgtgt	ttggggagaac	caggaaatca	tgtgtgggg	tcaatgttga	73320
agctgttgg	ctgggggtt	gggtggatat	aaggctggcc	ctggggagtt	tttccctgtt	73380
gaggccctt	acccacaact	caagatccag	tgctatagca	ggagatccca	gagctagtcc	73440
taacagatgg	tcaggattga	acttggccta	gagtaaaaatg	aggaggatg	tgccagaact	73500
ttctcaacat	acttttttgg	aagaggtcag	aaggcttaaa	gaggtatgt	aactggaaaag	73560
gggttcttgc	ccagacccca	ggaggggtt	tttgcaccc	gtcaatggaa	gagttcgaga	73620
cgagtcacc	cagtaaaatg	aaagcaattt	tattaaagaa	gaaacagaaa	aatggctact	73680
ccataagagca	gcgcacatgg	ctgcttaact	gagtgttctt	atgattattt	cttgattcta	73740
tgctaaacaa	agggtggatt	attttgtgagg	tttccaggaa	agggggcagg	atttccccaga	73800
actgatggat	ccccccactt	ttagaccata	tagatgtact	tcctgaccc	gcatggcg	73860
ttgttaactg	tcatggccct	ggagggat	tcttttagca	tgttaatgt	tataatgt	73920
tataatgagc	agtggggac	gccaggaggc	gttttcatca	ccatcttgg	tttgggtgggt	73980
tttggccggc	ttctttatca	catcctgtt	tatggcagg	gtctttatga	cctataactt	74040
ctccctggca	cctccatatc	cctcctgtga	ctaagaatgc	agcttagcag	gtctcagcc	74100
cattttacca	tggagtcgt	ctgatccaa	tgcctctgac	agcaggaaatg	tttggaaattga	74160
attactatgc	aagacctgag	aagccattgg	aggacacagc	cttcattttag	acactggcat	74220
ctgtgacagg	ctgggggtt	gttaattgtt	gttggccagt	gtggactgt	ggagatgtca	74280

ctactgttaag	atatgacaag	gtttcttcc	aaacaggctg	atccgcttct	tattctctaa	74340
ttccaagtac	caccccccgc	ctttcttcc	ctttccttc	tttctgatt	tactacatgc	74400
ccaggcatgc	tacgccccca	gctcacattc	cttccttat	ttaaaaatgg	actggggctg	74460
ggcgcggtgg	ctcatgcctg	taatcccagc	actttggag	gccgaggcgg	gcggatcatg	74520
aggtcaggag	atcgagacca	tcctggctaa	cacggtgaaa	ccccgtctct	actaaaaatg	74580
caaaaacatt	agccaggcgt	ggttcaggt	gcctgcagtc	ccagcggctc	aggaggctga	74640
ggcaggagaa	ttggcgtgaac	ctgggaggt	gaggttgc当地	tgagccgaga	tttgccact	74700
gcactccgc	ctggggcaca	gagcgcagat	ccgtctcaa	aaaaaaaaaa	aaaaaaaaaa	74760
tagctggcgc	ttgtggcgc当地	tgcctgtat	accagctact	ctggaggctg	aggaagaga	74820
atcgcttga	cccaggtaggc	ggaagggtca	gtgagccgag	atcttgcacac	tgcaactccag	74880
cctgggtaca	gagtggagact	ctgtctcaa	aaaaaaaaaa	agaaaaaaaaa	agacagaaaag	74940
aaagagcaca	gacagagtca	cagggttgc当地	cagttagga	ctgtcagggt	agagtgcacg	75000
gaaatagaaa	gtatattttt	cacttacagc	acatcttcgt	ttgatttagc	acattttaaa	75060
tactaatag	caacgtgtgg	ctttagtgc当地	ttcactaaa	tcttggacag	tgcaagtcta	75120
aagaatcctt	gatccgtccg	gcatggtgc当地	tcacgcctt	aatccagca	cttgggagg	75180
ccaaagggtga	aggatcactt	aaggtcagga	gttcgagacc	agcctggcca	acatggtga	75240
acctcgtctc	tactaataat	acaaaaaaaaa	ttagccggc当地	atgggtggc	atgcctgtaa	75300
tcccaaggatc	ttggggaggct	gaggcaggag	aatagcttgc当地	atccaggagg	cgctgcagtg	75360
agccagatc	atgcccattgc当地	actactgc当地	tccagcctgg	gcaacagagt	gagactgtct	75420
aaaaaaaaaa	aaaaaaaaattt	ttggggcgtgg	tggttcacgc当地	ctgtatccc当地	agactttgg	75480
gaggctgagg	gggggtggatc	acctgggtc当地	tgaggttgc当地	gaccagctg	gccaacatgg	75540
tgaaacccca	tctctactaa	aaatacaaaa	attagcttgg当地	cgtgggtgt	ggcacctgaa	75600
atctcagcta	ctcaggaggc当地	tgaggcaggag	gaattttctt当地	aacccaggag	gcagagggtt	75660
cagttagcgc当地	agatcgc当地	tctgcactc当地	atccctgggt当地	gcagagcaag	actatgtctc	75720
aaaaaaaaaa	aaaaaaaaatac当地	ttgatttgc当地	ggacattctg当地	cagaacatca当地	tatggagaca	75780
ctatgttgc当地	gacatcatgc当地	tgatttgc当地	caagaaatgg当地	caagtgttcc当地	agaaacacag	75840
tcaagacaca	taatgtccag当地	aagggtgagat当地	ataaaactcta当地	ctaagattca当地	gtggcctgccc当地	75900
acatctggta	cattttttaaa	cctgtctat当地	gtttgtgtat当地	aaaaggattt当地	aaccttggccc当地	75960
aaagaggggg	ctggggctt当地	tcccacta当地	ctggacataaa当地	tctttttaaa当地	ctcttgaat当地	76020
atcatttctt当地	atagaagttat当地	ttttttttt当地	actaggggcc当地	ttggggccagc当地	cagatagcaa当地	76080
caatgttgc当地	ttgggttgggg当地	gttttggatc当地	aggtggcatc当地	agtgtgacat当地	ccttgagtggc当地	76140
tagagacttag	aatcaaccac当地	atgggcagac当地	aacccagctt当地	acatgtatgg当地	attccaataaa当地	76200
agactttggta	cacaagggtt当地	tgggttgc当地	ttctctgggt当地	gcaatgtctat当地	atactgggaa当地	76260
accatttctg当地	atcccatagg当地	gagaggacaa当地	ctggatattt当地	tcattttgtt当地	cctccctggg当地	76320
ctttggccctt当地	tgcattttt当地	ccttgc当地	tttattattt当地	tattatgaga当地	tggaatctcg当地	76380
ctctgttccacc当地	caggcgtggag当地	tgcagtggaa当地	tgatctcaac当地	tcactgcaac当地	ctctgcttcc当地	76440
ccgggttcaag当地	cgattttctt当地	gtctcggct当地	cccgagtagc当地	tgggactaca当地	gatgcataacc当地	76500
accacaccccg当地	gtaattttt当地	ttgttatttt当地	attagagacg当地	gggttccatc当地	ttagccagga当地	76560
tgggtctcgat当地	ctctctgatc当地	catgttccgc当地	ctgcctcggc当地	ctctcaaaat当地	gcttaggaata当地	76620
catgtgtgag当地	ccacccgc当地	cagccccctt当地	ggctgattat当地	taaagtgtat当地	ccttgagctg当地	76680
tagtaaattta	taaccgtgaa当地	tataacagct当地	tttagtgc当地	tttgc当地	cttctagcaa当地	76740
attatcaaacc当地	ctaaggatag当地	ccttggggac当地	ccctgaactt当地	gcagtttgc当地	tcagaaataaa当地	76800
gggtgtctcat当地	gtgtgtacca当地	tgcctctaa当地	ttttgtat当地	aattaactt当地	cacaacttta当地	76860
tttattaccgc当地	ttacacttca当地	tttttgc当地	cattttatcca当地	cataccatt当地	attcttagtgc当地	76920
tttgc当地	agactttctt当地	tctcatgtc当地	tttatttgc当地	ttgaagtaaa当地	tccttttagga当地	76980
tattttttt当地	ttttttttaaa当地	ctttgc当地	acatacttt当地	atttttt当地	tattttaat当地	77040
tttgc当地	ttttttttt当地	ttttttt当地	ttttttt当地	ttttttt当地	ttttttt当地	77100
gatacaggca当地	tgcaatgtga当地	aataagcaca当地	tcatggagaa当地	ttgggtatcc当地	atccctctaa当地	77160
gcaatttatac当地	tttcaatgtt当地	caaacaatcc当地	aatattactc当地	tttaagttat当地	ttttttttt当地	77220
acattttaatt当地	tttgttgc当地	taggttgc当地	ctgttgc当地	atccaaat当地	ttttttttt当地	77280
ttttttttt当地	agagtctc当地	tcagttggccc当地	agactgaaag当地	tgcaatgtt当地	caagctcgcc当地	77340
tcacttcaat当地	ctctgc当地	ctgggttcaag当地	cgaatctt当地	gcctcagat当地	cccacatagc当地	77400
tgggattaca当地	ggcacacacc当地	accatgc当地	gctaattttt当地	atattttt当地	agttagagacg当地	77460
ggttttcgcc当地	atgttggcc当地	ggctgtt当地	gaactctt当地	cctccaaat当地	tctgaccacc当地	77520
tcagccctcc当地	aaagtgtat当地	gattacaggc当地	atgagccacc当地	acacccctggc当地	aaaatagaat当地	77580
attctttagt当地	gagggtctgc当地	ggtgacaatt当地	ttttttttt当地	ttttttt当地	ttttttt当地	77640
gttgc当地	ttgggttggag当地	tgcaatagca当地	cgatctc当地	tcactgcaac当地	ctccacccctt当地	77700
cggttccag当地	caatttctt当地	gcctc当地	cccaatgtc当地	tgagagat当地	caggcaccacc当地	77760
ccaccacacg当地	cggtcaat当地	ttgttattt当地	attagaaatg当地	ggggttt当地	ttgttggcc当地	77820
ggctgggtt当地	gaacttctg当地	cctcaggtg当地	tccacccacc当地	tttgc当地	aaagtgtctgg当地	77880

38/122

gattacaagg	atgagccacc	acgcacagcc	aatttttcc	gtttttgtct	gaatcttat	77940
tttggcat	cttggaaata	tatTTTgt	ggatataaaa	tttgtgttg	atagttatta	78000
tcattattat	tattatttg	agacagggtc	tcactctgtt	gcctatgtg	gggtgttagta	78060
atgtgatctc	ggttcaactgc	agacttggacc	tcctagggtc	caggtgatct	tcccacaccta	78120
gcctccctag	tagctggac	tacagatgca	tgccaccata	cccaactaat	tttcttattt	78180
ttttagaga	tgaggctttg	ccacattcc	caggctggtc	tctaactcct	gagctctagc	78240
aatccaccca	ccttggccat	acaaagtgt	gggcccgtac	tagccagcag	ttactttttta	78300
tagcatattg	aattattat	atgaatcttc	tgcacccac	tgttaactgtt	aaaaaaatca	78360
gctgttact	tggcaactctt	ttttttttt	tttttttga	gacagagtct	tgccctgtcg	78420
cccaggctgg	agtgcagtgg	cgtgatcttg	gctcaactgca	agctctgcct	ccccgggtca	78480
cgcattctc	ctgcctcagc	ctccggagta	gctgggacta	aaggccccc	ccaccacgccc	78540
cggctgattt	ttttgtattt	ttcgttagat	ttgggtttca	ccgtgttagc	caggatggtc	78600
tcgatctctt	gacccctgtga	tctgtccgc	tcggcctccc	aaagtgtgg	gattataggc	78660
tgagccacc	gcgcggcagcc	tctttttttt	ttttttttag	acggagttt	actctgtcat	78720
ctaggctgg	gtacagtggc	gtgatcttag	ctcaagtgc	cctccacctc	ctgcctcagc	78780
ctgccaaata	gctgggattt	caggtgcgt	ccatcacgcc	cggctaattt	ttgtattttt	78840
agtagagatg	gggttttcc	atgttagaca	ggctggcttc	gaactctgg	cctcaagtga	78900
tctgcctgcc	ccagccctccc	aaagattaca	ggcatgagcc	accgcaccc	gcacaaggtagc	78960
actcccttga	aggtatctgt	cttcccctac	ccctagcaat	ttttaacaat	tttcttcat	79020
ttttatattcc	tgaagttttt	ttattaataa	tctgtgtgc	gattttttt	tattttttt	79080
gtttgcagtt	catagtgatt	tttgaattt	tgtgttggtt	tctgttatca	ccacaggaaaa	79140
atttgcagcc	tttagctttt	caaataattt	tttgcataat	tctctcttct	ccccccttccg	79200
tacaattat	tttggataaaa	ctaaaacttac	ggccgggggtc	agtgactcat	gcctgtatc	79260
ccaaacactt	gagaggctga	ggcagggtga	tcacctaagc	tcaggagttc	aagaccagcc	79320
tggccaaat	ggtgaaacc	cgtctctact	aaaaatacaa	aaattaccag	gcatggtggc	79380
acacatttgc	agtcaggagg	ctgaggcagg	agaatttgc	gaatccagga	ggtgagggtt	79440
gcagtgagct	gagatccac	cactgcagtc	ttgcctggc	gacagagtga	gatggaaatc	79500
tgtctcgaaa	aaaaaaatgtt	tgtattttgc	ataaaactata	tttggtagaa	tgtttgttgc	79560
agaataatcat	tcattgattt	ttaaacaat	tttggataaa	ccattcaactg	gatttgcgtat	79620
aattaactt	ctgatTTTAC	ctcaactgatt	tgttgcattt	aatacaactg	gtataaaaaag	79680
actgtgacga	ggccgggcat	gggtgtctcc	gcctataatc	ccagcaactt	gggagggtga	79740
ggcaggcgg	tcacctgagg	tcaggagttc	aagaccagcc	tgaccacat	ggtggaaaccc	79800
catcttact	aaaaataacaa	aatttagccgg	tcgtgggttgc	gcatgcctgt	aatcccagct	79860
cttcgggagg	ctgtggcagg	agaatcaactt	gaacccggga	ggtggagggtt	gcagtgagcc	79920
gatatcgcc	cattggactc	cagcctggc	aacaagagcg	aaactccgtc	aaaaaaaaaaa	79980
aaagaaaaaa	aacacataaa	acaaaacaac	actgtgacgg	ttcccaaaaa	ttaggagcat	80040
aattaatgg	actcctgata	aaaattttaa	ttatcttaca	tgtaaaacta	aatgacttta	80100
tgaagttat	ttagaaaaat	aatgcagggt	attagttgc	cacagctgc	tattcagcct	80160
aatgtatatt	tcttgcattt	tttaaattct	tcttttact	ttactcatat	gtggatcatc	80220
aaatttcaaa	agattaaatg	acaataactct	tagcagcaag	tttccctaag	catataaaaca	80280
ttttatatgg	tgtatgatttca	gaaggatccc	gaagaatatg	tactgcaga	tatcattcac	80340
ccccatatac	ctgcccggaca	gacatccat	tttgggaccc	tggataatg	tgtgggtgg	80400
gagaaagatg	ggagaaaaatg	gtataagcaa	atggctttgg	agtctgttgc	acagcgatttgc	80460
aaatccctgtc	tctaccttct	aacagccctca	tgatcctaca	taagttaccc	cgatccctcag	80520
ggccacatct	gtaaatttgg	ggttgcgtat	gcagccatct	cacagggtct	cttttccgggg	80580
aagggcagga	attatggatt	aagttagtca	gtatgttgc	agcacttaat	acaaggagggg	80640
cgcataataa	gtacttcata	aataatgtac	gcattatca	tgactgttgc	gtatgcagct	80700
gtcggggatt	acggcgactt	cagaattttt	tttggggcagg	gctcaaggc	agcaaatcac	80760
actggaaatgc	gaggtggaggc	actgttctg	cacagactgc	ttagctggag	agaatgagga	80820
aggcttagag	gagattttgc	ggaacttgc	tttgcctccgc	tccaaactctg	ttggatctgc	80880
ccccctgcca	gagacattca	ggggattttt	cgcaactctcc	cctcccttac	gtccctcccg	80940
ccccatccaa	ctaaccacac	aacacataca	aaatagcccc	tgcgagggtt	tgcaegctgg	81000
aaggaaacag	gagaaggggcg	ctgcgttcc	tttgcgtatgc	cctgtacttgc	ggcccccgttgc	81060
agacacagcc	acttgcctcc	tcagcgtca	gagaatatccc	acgttagaccg	cgccccgggtc	81120
cttgcgttca	gccaatctcc	ttttgggtgg	tttggggatgc	acgatccaag	gttttatttgc	81180
ctacagacag	ccccgggttgc	tccgccaaga	acacagatttgc	gttcccggagg	gcatctcgaa	81240
ccccctgggt	ggccggccgtc	agcctcccg	tgcaggcccc	gccgaggcc	ggaggaagcg	81300
gccagaccgc	gtccatcg	cgccagctca	ctccggacgt	ccggaggctc	tgccagcgct	81360
gcttccgtcc	agtgcgcctg	gacgcgttgc	cettaacttgc	agaaaaggctt	cacccgttgc	81420
tccaggcttc	atccctagt	agcgtgtgc	tttgcgttgc	tgacttttattt	tttgcgttgc	81480

39/122

tagtttcca	gataccagga	ctgactccaa	ggactattac	tcatctggag	ggtttagcac	81540
agtaccgtcg	catagtaaat	ttccatgtca	gttttggta	cctttcatgc	acttgcaaac	81600
atgccatgct	ctgaaacgaa	ataggcacat	ctttttttt	ttttttttta	aggagtcttc	81660
ctctcgccca	ggctggagtg	cagtggcgcg	atcttggctc	actgcaacct	ccacctcccg	81720
tgttcgagat	tctcctgcct	cagcctctg	attagctggg	actacaggca	tgccacgacg	81780
cccagttaat	ttttgtattt	tttagtagaga	cggggttcg	ccatcttggc	caggctggtc	81840
taactcctga	cctcaggtga	tctgactgcc	tcagcctctc	aaagtgttgg	gattacaggc	81900
ataagccact	gcatctggcc	agaaaatgaaa	taagtaaatc	ttaaaatctg	ctctaacaat	81960
atagtaaaaa	gaccatatta	ttattagac	agttaaaggg	atttgccat	ttcgggttct	82020
agttatagtc	ttaaaatgg	acattctgt	agaaaataaa	aagtttctc	ttcaaagttc	82080
cccttcttgt	taaagaatac	atcataagtg	ttagaaagtaa	tagtttattt	taaagactaa	82140
ctttcttcaa	gcctccttgc	tttgcgttaa	taactctttg	ttaaggccct	tcctatgtaa	82200
ctgttggaca	tgctcacagg	cacgttccag	ttcacagctc	atgcccccttc	cttatttgga	82260
aatgttattg	cttcctttaaa	cctttcggtt	agcaacttcc	tctccttctt	cgttcttct	82320
tgcaacttacc	tatggaaaaa	gttttaggtt	attagcaaat	cggctatcag	tttaagagtg	82380
tgaggtcccg	ctccagccaa	tggatgcagg	acatagcagt	gaggacgacc	caaatgcgtt	82440
agggataaat	atgtttgctt	ttccttgcgtt	caggtgtgt	ctcgacatcg	ttccatctgc	82500
gattgagcac	cctttctgca	gaaagtaaag	attgccttgc	tggagatctt	ttgtctccgt	82560
gctgactttt	cttcgtggca	ccgattatct	atttctaaca	attttgttat	ttctaacatt	82620
ctgacaatac	ttgggttagt	tgtcttctt	gggcctgttt	ccccatccgt	cacatgataa	82680
acttcattgg	ttttaaaaacc	ccagcgaaca	tttattgagt	tactattacc	ttctgcccct	82740
ccccaaacccc	aaccccgagg	agcagttaca	acctcagccg	ctgagcgcac	tcgcccgggt	82800
ttaagaagca	ccaaagacag	ggaggcttga	tttgcatttgc	tttggggata	gagggtcaga	82860
agatccacag	gaaaatggca	tttgcgttgc	gttgcgttgc	tggagcttc	ttttaataac	82920
tggcggaggt	tttatgttgc	agtccttac	aaagttgagc	attcgcaagg	actgcactcc	82980
gaaataagcc	cgcttccctt	tttcatttgc	taatgatcca	gggagctgt	ggttccgcac	83040
gcccgggggtt	gtgccttttc	ctaattcagg	ttctgcattcg	cctcgaaccc	gcaggccgt	83100
gccccgttctc	ctgaggaagc	agggacttgg	gtgcagggttgc	aagctgtcg	tgcggccag	83160
cgccctgtgag	caaaaactcaa	acggaggaggc	aggagggggtc	gagctggagc	gtggcagggt	83220
tgaccctgcc	tttttaggg	gcacaatttgc	aaagggttacc	agggggccgg	agccggggac	83280
ctaaggcccc	ccccgttcca	gctgtggga	gggctcccgc	cccagggagt	tagtttgc	83340
gagactgggt	ctgcagcgt	ccacccgggg	ccggcgcacag	acgcccacaaa	acagctgcag	83400
gaacgggtggc	tcgctccagg	caccccgaggc	ccggggaaaga	ggcgcgggt	gcacgcgcgg	83460
gtcacgtggg	cgatgcgggc	gtgcggccct	gcacccgcgg	gagggggatg	ggggaaaagg	83520
gccccggccgg	cgcttgaccc	ccctgttgc	ctagcgcggg	gaaggaccgg	aactccggc	83580
ggccggcttg	ttgataat	ggcggcttgc	gtgccttgg	catccccagg	aggcgggtgg	83640
gcccaactccc	ggaagaagg	ttccctttcg	cgctagtgc	ggccggccct	tggaccggaa	83700
agtccggggcc	gggtgctgaa	tgaggggagc	ccggggccctc	ccgcgcgcgt	ccccccgcac	83760
cctccgtcccc	gacccggggc	ccgcattgc	tttcttccgg	cgaaaggtt	gctgaggggg	83820
cgccgggggg	gagtccagg	gggcctcagg	ggcggcggttgc	gggcagggttgc	gcctgcgagg	83880
gctttcccca	aggcggcagc	aaggcctca	gcgagcctcg	acctcggcgc	agatgcccccc	83940
tgagtgcctt	gctctgtccc	gggacttctc	ttgggagggag	aagggttgc	tcttgcgcga	84000
ggtcagagga	gtattgtc	gctgggttc	aaagcgttgc	taaaggccat	agaagttcct	84060
gcctgtttgg	ttttaaaaac	ttctttaggt	gggggttagt	ttttgttgc	tctttgagga	84120
ccgtggatca	agatcaaga	aatctttta	gaaccttatt	atggaagtct	gaagtttcca	84180
aatgttgagg	gttttatgtc	taaaagcaac	acgtaaaaaa	attgtttct	tcacccagt	84240
ctgtcttcca	atttccttctt	ttggggggagg	ggtagttact	gtcttacta	aaataaaaatt	84300
acttattgtct	aaagttccccc	aacaggaaaga	ccactactt	tgatgactt	gcaagtttgc	84360
ctaactactg	gaacccttaac	ttacaaacga	actacttaca	ttttgttgc	ccagtgttat	84420
tacccgttcca	atgtttactgt	agaaaacagct	taattttgc	tctgggttac	gttggcgcac	84480
tccattaaaa	atacatatcc	gaagtggagca	agtatgggtc	tgtggacagc	agtgattttt	84540
cctgtcaatt	cctgttgc	cagataaaaat	gtaccagaca	gaggccgggc	gcgggtggctc	84600
acgcctgtaa	tcccagact	ttgggggggt	ttgggggttgc	atcacctgag	atcggggagtt	84660
caagaccagc	ctgaccaaca	ttggaaaacc	ccgtgttac	taaaaataca	aaatttagcca	84720
gggtgggtggc	gcatggctgt	aatgccagct	acttggggagg	ctgaaggcagg	agaatgcgtt	84780
gaacctggga	ggcgggggtt	gcccgtgagcc	gagatagc	cattgcactc	cagcctgggc	84840
aaaaagagcg	aaactccgtc	tcaaaaaaaaaa	agtaccagac	agaaatgggt	tttgggtttct	84900
ttttttgttt	tgagacggag	tttcgttctt	gttgcggccagg	ctcgagtgc	atggcgcgtat	84960
ctcagtcgtc	gctcaactgca	acctctgtct	cccagggttta	atcgatttctc	ctgcctcagc	85020
ctcccaagta	gctgggat	cccatgcccc	accatgcccc	gctaattttt	gtatgttttag	85080

tagaaacggg	gttcaccat	gttaggctgg	tcttgaaccc	ctgaccta	gtgggcctcc	85140
cacctcgcc	tcccaaagtg	ccaggattac	aggcatgagc	caccgcgcc	agccagaaaat	85200
gggttttgg	aaaagca	aacaaaatcg	aacttggttt	catatgacag	ctctgctgct	85260
aactgtaca	ggggcagacc	agttaa	ctttctgtc	ttctgtcagc	tgagaattag	85320
atgatcccc	aaggcccatt	gaactctgaa	tgactttaaa	tacttcttct	taagtggta	85380
cacggttt	gtactgtat	ccaggtat	aatgcataa	agtgccta	aatgaaaacc	85440
gtaaaatag	taggaggaag	cttatttgtt	aaggcagggg	tataccta	agcttctaa	85500
tttatggta	ttgaagtgg	taactttgtt	ttttttttaagg	ggggaaaca	ttctaagaat	85560
aataggcca	actgcataat	gcacaagaga	ctgtgtctc	tattca	atacctttt	85620
agtgtccaga	gtctccagg	tgctgtcta	ggccctcacg	attgatgt	gaaccagaga	85680
atgtccctgc	acccatggag	cttatttgtt	actggggtag	acagataata	aataagcaaa	85740
caaatttct	cttttctccc	tttcgtcca	tgtta	tgtgtatagg	tgtataactt	85800
caagttgagt	aaagtgttat	gaaagattaa	gaggagaaat	gcattttgt	tagatgttag	85860
aggactcagc	aggtgaccc	gaaacttaga	gctgaaggat	cagttaggg	taactagaga	85920
ggccagggaa	tcgcataat	aaaggccagg	aggcagaaaa	gagcatgtg	cccttcaaga	85980
gaggaaagaa	ggctactgtt	actggagcat	agatgtaggc	aagtgttgg	tgattgagag	86040
ctctacgggc	catggtagg	tttatttct	aatgcgcgaga	tgccaaacat	gggtgttcat	86100
atctgtata	ccagtatttt	aggaggccga	ggcaggaaat	tagcttgaac	ccaggagttc	86160
aagccagcc	tgagcaacat	gagacatgt	caaaacattt	aaaaaaatgc	tggtatgtat	86220
gtgcacacc	tgtgttccca	gctactcagg	aggctgaggc	agaaggatca	cttgagccta	86280
ggagggtgg	gctacaatga	gccatattt	agtca	ctccagcctg	gatgacaaag	86340
tgagaccatg	tgtcaaa	aatacagaaa	gaatattaat	ttaaaattt	gaaagaggag	86400
tgatctgaac	ttatatctt	aaaagatcat	tcttagggat	ggtggtccat	gcgtgtatc	86460
aagggtttt	ggaggctgg	acaggaggat	cac	cagttaggaa	tcaacctgt	86520
cagcatagag	agactccat	tcttacaaaa	gaaaaaaat	atagctgg	tgtgtgagtt	86580
attcaggagg	ctgaagcaga	aa	gagcccagg	gtttagg	gcagtaagct	86640
atgatccac	cactgcaaca	cagttagat	ttgtctt	aaaaaaaaa	aatcattcta	86700
ggtgtttt	ggaggctgg	tgtgtt	gtagaa	gagatgttcc	tgttagggat	86760
tcgattcaga	ctttaataat	catcaatgc	tttagtccca	aatttacatc	actacgttgg	86820
atccttgc	ctgaatccat	actggat	ttcaactttag	gtttagttt	tatcttacc	86880
tgaccat	agagggtgtc	agtcttttgg	cttccctt	ccacatttgg	agaagaattt	86940
tcttgagcc	cacatagat	acactaacgc	taacaatagc	agatgagct	aaaaaaaaatc	87000
gcaaaactt	taatgtttt	agaaagtta	cgaaatttgg	ttgggcacat	tcaagccat	87060
cctggccgc	gggatggca	agcttaatcc	agtagatacc	ttcaacttac	aatatctaaa	87120
atttatgg	agatttatgc	atttttaacc	tgttcatc	tttttcttca	gaagttagt	87180
tttggctt	tttcttttct	tttttttgg	atggagttt	gcttttatcg	tcaagctgg	87240
agtgcagtgg	cgatcttgg	ctca	cttccgcctc	ctgggttca	gtgattctcc	87300
tgccctcagcc	tcgcaagtag	cttca	aggatgcgc	caccatgacc	agctaatttt	87360
tggagacagg	gttccatc	gttgg	tttgg	actctgtacc	tcagggtatc	87420
tgcctgc	ggcccttccaa	aggctggat	tacaggcat	agccacccgt	ccccgctgca	87480
tttttggatt	tttagtgc	cagccaaaa	ctttagtaca	tctttgaacc	tcttttttcc	87540
tcctactcta	tatctgtatcc	atcaga	ctttaggtc	tacctcacac	atatcgaaat	87600
cctaccacgt	ctcaccatct	gtgacat	acaccctgg	ctaggcgtc	atctctgtt	87660
agattgtgt	gtttaggtat	ttcttcaagg	agatgacat	caaatttt	cttaaatgtc	87720
aagaggggagc	ttttttata	aa	gatggagg	aggcagatt	attttgc	87780
tttggtttca	tttcatttctt	gatactt	tgatata	aaaacaaat	cacagaaaca	87840
gaccaggta	tatttggaa	ttcgatata	gatgttctag	ttggggaaag	atagactgt	87900
ctgttaatga	tgcttagtt	ccatcatctg	gcaaaaaata	at	ttctctcata	87960
tatctcagat	caacagactt	tttctgtt	ggggccaaatc	ataaaat	tagcttcc	88020
agaccatgt	ttttctgtca	actcttctt	tatcctt	gcccata	atatgtaaac	88080
aaatggccat	ggctgtgtca	cgataaaact	ttacttaca	aaactgtt	ttggccagtt	88140
taggcattgc	cagca	ggaggat	gcatgtt	cacttgg	caggagttt	88200
agaccaggct	ggccaaacat	gtgaa	accct	aaaat	aaatagctgg	88260
gcatgttgg	gggtgtctat	aattccat	actctgg	ctaagacaca	agaatca	88320
gaaccaggaa	ggcagagggtt	gcagt	gagatgac	cactgc	caggcagggt	88380
gacggagtct	taaagaaaa	caaaacaaa	ggtagtgg	tgtatttgc	ccatgggt	88440
tagttgcca	atccctgtat	caga	ttccaggta	ataagac	ggaatgtt	88500
aaaaacaaaa	cttgaagtca	tgt	gaa	caatcc	gatctcagga	88560
taggaaggga	tattgtttaa	aataagacac	aggaaaat	aatccat	gtgtaaattt	88620
gactacgtt	aaactaaaa	cttgc	ccaa	gcccgg	tcacgcctgt	88680

ctttgggagg	ccgaggtgag	cagatcacca	ggtcaggaga	ttgagaccat	cctggctaac	88740
acggtaaac	cccgtctcta	ctaaaaatac	aaaacattag	ccgggcgtgg	tggcgccgc	88800
ctgttagtccc	agctacttgg	gaggctgagg	caggagaatg	gcctgaacc	gggaggcgaa	88860
gcttcgtcgt	agctgagatc	gcccactgc	actccagcct	gggcgacaga	gtgagattcc	88920
gtctaaaaaa	aacaaaacaa	aacaagcaa	aaaacctaaa	actttcatac	aataaagtat	88980
acctaagata	cttctagaag	agaagattt	catccaggac	gtgtatggaa	tttctgcaag	89040
taataagtaa	aagacaaggg	acatgaagag	gcagttcaca	aaagaggaaag	ccaaaatgac	89100
caataaacat	gaaaggatgt	ttaacctca	agaaaacaag	gaaatgaatt	aaaacatca	89160
aatgccattt	caaaaactgt	aagttggcaa	aattaaaat	accaaggatg	agaatatgaa	89220
gcatggctat	atgagtgcat	ggaatggatc	agtcaacttc	attaaaaatg	cacataattt	89280
gtttttatt	tatTTTTT	agacagtcta	tgtcggccag	gctagaatgc	agtggcatga	89340
tctcgctca	ccacaatctc	tgccttctgg	gttcaagcaa	tttccctgc	tcagcctcct	89400
gagtagtgg	gattacagge	acatggccaa	acggccgggt	aagttttgtt	tttttagtag	89460
agacagggtt	ttggcatgtt	ggccaggctg	gtctcgaaact	cttgacctca	ggtgagctgc	89520
ttcccaaagg	gctgggattt	gaggcgttag	ccaatgctcc	tggctgaaaaa	aaatgcacat	89580
aatttgttac	ctagcaattc	catgtctaga	ggcttatcct	agagaattt	ttgttataat	89640
gcataaggaag	acgtgtacta	gaatgttac	tagttgaatg	tttaagtgaa	aatttaggaaa	89700
taaagtaataa	gttcaataac	agggaaaatg	gttaaaggat	atttataaaa	caattaagta	89760
gctaaaatgt	ataaaactaga	gctgcgtgaa	tgaactgaaa	ctggttcaat	agtcatgtca	89820
gattttgtt	taaatacagg	tcagatatgt	atagagtgtc	atttgcgtt	ttaaattttt	89880
ttttttttt	gagatggagt	ctcactctgt	tgcccaggct	ggagtgcagt	ggcgtgatct	89940
cagctcaactg	caacccctcac	ctcctgggtt	aaagtgttcc	tcctgcctca	gcctcccgag	90000
tagttggat	tacaggcatg	caccacatg	cccagctcat	tttccctattt	ttatggccca	90060
cagggttca	ccatgttgc	caggctgtc	ttgaactcct	gacctcaagt	gttccaccca	90120
acttggcctc	ccaaagtgt	aggattacag	gctgtgagcca	ccgtgtcag	ccatttgcgt	90180
gattttttaa	gatgtgcaga	ataatgccc	taaaaaaaaaat	acacatacat	gtatataat	90240
acacgtttgg	ctgggtgtt	tggctcacac	ctgtatccc	agcaacttgg	gaggctgagg	90300
caggaggatc	acttgagccc	aggtgtacaa	gactagccctg	ggcgagatag	caagacccca	90360
tctcaacaac	agaaaaggata	attaggtatg	gtggcatgag	aggatctact	gagccccagga	90420
gttgcgttgt	tatcaggcct	ctgcaactct	gcctggacaa	caaagcaaga	ccgtgtctca	90480
aaaaaataaa	aataaaaaat	atttgcgtgt	ggtcatagtc	aaaaaaacgt	catggaaagga	90540
aaatgtcttt	atttatttat	ttatTTTTT	ttttaaga	cagacttgg	ctctgtcacc	90600
caggctgggg	tacagggtt	taatctcag	tcaccgcaat	ctcggcttcc	cgggttcaag	90660
cgattcttct	gcctcagct	tctaaatgtc	tgggactaca	ggtaccggcc	accacaccct	90720
gctaatttctt	gtgttttcag	tagagacagg	gtttcaccat	gttggcaagg	ctggctctcga	90780
actccgtacc	ttaagtggac	cacccgcctt	ggcctcccaa	agtcctggga	ttacagggt	90840
gagccactgc	gcttggccag	gaaatatcta	atttgcgtt	tggaaaggaa	90900	
agggtcaggt	gggttccat	aggaactcta	aagtctat	ataatactta	gggggacaga	90960
agggaaataaa	gcaaaatgt	gatattttgt	tgttgcgtt	tgtatatgtt	agaagtataa	91020
cataggatgt	ctgattgtata	gtaggagaat	gttttttaggt	ggtaaaatgt	gaaccgtggt	91080
ggtttggttt	ggcagtagaa	tcagttggtc	atagtttgc	tgtggaaagg	aataaaacaga	91140
ccatgttaag	gatgacttcc	ggaattttgg	tctgagtagt	gggtggatgt	cagtgtcatt	91200
catgagggaa	gatgaaatgt	gaggtaggaa	cagggtttgg	agaagatgt	atgttccctt	91260
ttagacaatgt	ggaattatgtt	aaagatggcag	gtaggtgtt	agctatata	atttgagata	91320
aaagattttag	gttggagata	taaattttgtt	agtaacagcg	tatctatgtt	attgttgc	91380
ttaagaatgg	gttaggtatc	ccaggaaata	cagatgtata	tgcagaagag	aggagtcaag	91440
gaagccaaga	caagttaatg	tttaaagtgt	gtgtatgtt	ccatggccag	atgtgtctga	91500
gagggtctca	aacccatgt	accctacaaac	atTTTTAAT	gtcgcttcc	tgacagcagt	91560
gatcgttacc	tgcacgtat	tttattttttt	tttcatgtt	agtctccaca	cacttgaatg	91620
tagacttttt	gaaggcaaaa	tcattgcctt	ttctgagctg	ggagcatgtc	tggcacatac	91680
caagcactca	acagttgtat	tattgtactt	atccagatac	tctgaggggcg	agttattttcc	91740
tgctacttagc	cttgcaccc	tcaatgttt	agaggcaca	tacagatgt	ggcacgtttt	91800
ggcatttttt	atTTTTGATAA	cctttccctg	gttgcgtttt	ttaatgtt	aaaaaaaaaaa	91860
caagaaaaaa	gggtttaaaaa	tagttttat	tcagatctt	tgtatgtt	cacacttggc	91920
ttaagctgtt	gggcaccc	ctatTTTGA	tttgcattttt	gtttatctac	agcagaattt	91980
ttactgtttt	atgttagtaag	gaagcaattt	tatgttattt	ttacagacaa	attattttttt	92040
atctttttt	tttttagacg	gagtttctt	tttgcctccca	ggctggatgt	cagtgtcg	92100
atctcggttc	actgcaccc	ccgcctccctg	gttcaagca	attctctgcc	tcagcctccc	92160
aagtagctgg	gcttacaggt	gtcccccacc	acaccccgat	cattttttt	tatTTTTAGT	92220
agagatgggg	tttcaccatg	ttggccagge	ttgttcttgc	ctactgaccc	cagggtatcc	92280

acccgccttg	gcatcccaa	gtgctggaat	tacaggcgtg	agccacccgt	cctggcccag	92340
acaaattt	atactctgag	tgtagaggc	tttagatgtt	ttcaactgtat	gctatggag	92400
gaataagtaa	taagatatg	tacacaacca	aagacccccc	ttcaactatgc	ttcttagatgc	92460
tagtactatg	gatgacacat	ggtaataata	ttggtagca	tttgcctca	attactgtg	92520
ctagttactc	ttctaagccc	cttacaggt	tatattttt	ttcatcaata	atccctctaag	92580
gtagtttta	ttattgacct	aatttataa	atcaagaaaa	ttaagaccca	gagaagttaag	92640
taatttgtcc	aagatcacat	ggcttataag	ttgtagagcc	agaatttgac	cccgatgtt	92700
gtgactacat	tgtctctca	taagcagggt	caacttttt	gactggatgc	tgttccaagg	92760
tcacttcctt	agagaagct	ttgctgacaa	ctaccctctt	gtgcctcct	ccaaggctgt	92820
ccattgttct	agaactttga	atactcatct	tagaataaaag	ctggctaat	tttacagt	92880
ttatagaatg	gatctctgac	tgcaaaagg	ggtcataatt	atcttttat	gttcttagt	92940
aaggcaaaga	acaagagaag	acctcagat	tgaagtccat	taaaggttaag	ttctgccctt	93000
ggcagtcac	tgcattaaaa	agtgtgtc	tttgcatttgc	tgaggctttt	aatctgtt	93060
tactctctct	tttggcat	atccatctg	ccttattttt	taattactta	tgttttgat	93120
ttatccctt	ctttaacctg	tataatgctt	taacatctag	catataataa	gtaggctttt	93180
ttttttttt	tttttttgg	gacggagtct	tgctctgtt	cccaggctgg	agtgcagtgg	93240
cgcgatcttg	gtcactgca	agctctgtt	cccggttca	caccatctc	ctgcctcagc	93300
ctccccagca	gtctggact	cagggtcagc	gcgcacgc	tggcttaattt	tttgttatttt	93360
tttagtagaga	cagatttca	ccatgttagc	cagtatggc	tcgatctctt	gaccctgtga	93420
tccggccgccc	tcggcctccc	aaagtgtctt	gattacaagg	gtgagccacc	gcacccggcc	93480
gtaagttaggc	tttttttacc	ttaattttat	ttttttgaga	tggagtttt	ctcttattccc	93540
caggctggag	tgcagtgggt	ccatctcggt	tcactgcagc	atccacccctt	cggttcaag	93600
cgatctctct	gcctcagcct	cccggatgt	ttggattaca	ggtggccggc	accatgccc	93660
gctaattttt	gtattttag	tagagcagg	gtttcaccgt	tttggccagg	ccagtctcaa	93720
actcctgacc	tcaagtgtatc	cactcgccctt	ggcctcccaa	agtccctgg	ttacaggcgt	93780
gagccaccat	gcctggccat	aaagttaggtt	ttactgagcc	ttgtgtgtat	tggtctatcct	93840
agtgattaca	gtgaaccagt	gccccttta	ttaatcacac	atthaattgt	tccctaaaag	93900
tgattatgtt	actttttttt	tttagtaaga	caaaaaaaat	agaataactt	taactgagca	93960
gtctgttac	tgttagggaa	cactgacact	tataaggctt	atgttttgc	cattttatcca	94020
gaagatgtt	tgattacagt	ttttttttt	ttatgtt	gaacaacctt	aattttaaat	94080
atattttgtt	tatTTTTTGT	ttggatcgt	acattgtctt	tgtttataga	tttagagcatg	94140
ctttttaaag	atgctgtatt	actactgt	tttattttgtc	cagtgtacag	agattgaagt	94200
gggaaaatta	taatggaaat	ttgtttccata	gtcattacat	atthaatttca	tcaattttt	94260
tccataaaaat	ctgttagattt	cttactttt	agatTTTCC	ttcaatgtt	tttatgtt	94320
attgcttgca	ctgagttttt	attctatatg	ctcaattttgc	ttggagaagaa	gactaattat	94380
aacttaggca	atggggggat	tggcttagct	atTTTCTAA	tctctactgg	gctaaaaaat	94440
atttctttag	taaagccagt	tagatccac	attagtttca	actgcctt	ttgagcaaaa	94500
cttgattggc	agtgtataaa	gcttaaaggcc	cttctcaagc	agagactgt	aaagactaga	94560
tctgtactgt	gtagaaggaa	ggaaacttaga	tttttcaggc	agtggaaaca	ccagtcttcc	94620
actctaaact	ttggccactaa	cagtatgacc	ttggggagtt	gtacttttct	tcagattttt	94680
catttggta	atggggggat	tggcttagct	atTTTCTAA	tctctactgg	gctaaaaaat	94740
tctgtgttta	tactctgtt	atgaagtaca	taatctgtc	ttaacattca	ctgactt	94800
cttaggataa	tacagaagca	gtacaagaaa	cagccctca	agatgttgc	agctgtggta	94860
gaaagacaaa	tttataacaca	gaacagttagc	aaatagacca	aaataataat	agtcgcatt	94920
tatagaacac	tttttctgtt	ctgggcat	gacaaaaact	gactataacg	gtgacaaaaa	94980
aagacttagg	ttctggccctc	attgaactta	cagatttagta	ggggagaggg	acattaaatca	95040
agtaattcca	cagatggcc	agcctagatt	ggtagtgtat	gaagtaaaga	gatgtgaacg	95100
gacttggaaa	aaaatttccgg	ggccaaaatgg	atagaagttt	attattttt	aaatatgagg	95160
tgttagagag	agggatattt	aaatgtata	cttacatctt	ggcttgcctt	acagaaccaa	95220
aacagaaaaat	tatatgttca	gtttttttt	tttgggttggg	aggtgtttt	gagtcttca	95280
tttatatatgt	ttatatatgt	tatTTTAT	gtatgtat	ttttaaggctt	gagttttaaa	95340
ccaaagggtt	gagagtgtt	tttttagatc	tagcaaacct	aagttgaaat	cctgcctgtt	95400
gaaatggctg	tttactatgt	cattaaacca	ggccaaaatgt	ttcaacttgc	tttcattttt	95460
gtcttcatct	ctaaatgtt	gaaaatatgg	tcttacaaga	tttgccttgc	agatagatga	95520
aataatatcc	aaaaaaaaaa	aaggatcata	gagaaactcg	tatagtgctt	ggtatatagt	95580
aggtcctcca	ttggtagcta	tcattatcta	gttttacat	agccttcagt	ttgttgaatt	95640
agtcaaactg	agtgtttttt	tgcaagggat	tcaaggat	ttgagatcaa	caaattttt	95700
ctgaagtttta	ggggaaagactt	catggcaatgt	acacttacct	tgtataaaag	ttgaaagaata	95760
agaaagatttt	gaatggagaga	ttttttctct	tctccctacc	agcccagctt	tttattttgag	95820
gatatttttttgg	gcaaaaggggc	cttcagacaa	gtagagggg	atTTTACAG	aaagattttgag	95880

atgaaggtagt agaaggctgt	aaagaccaga	aaagagaatt	gagacagagg	aagcaggaag	95940
ccactgttagg ttttgagca	agatattgtat	gctgttaagta	tggtgttat	gaaaggtag	96000
tctggaaagag atttgcaggta	tggagacccc	ggaagtttt	ttgttataat	acagaaagac	96060
ttgcaactgag ggtgagggtgt	taaaaataaa	caggttaagta	aatgtttaaa	catcttgaag	96120
gaaaagtcaa caaatcttgg	caagtaaaca	gataacagtg	aaaaagaatg	ggaccaagat	96180
tttgagttt ggagactgg	ggattgaaca	gacagggaaa	ttgagaggag	aatcagatga	96240
tgatgttta atgtgtat	tagacaggat	gtgcttgaga	tggttaaagt	aatgtgggtg	96300
ggaatgtcta gtacgagta	atcagtata	caagacaaa	gcccgagtca	aagacaagt	96360
acagatacag atcagggctt	ttcatctgc	tccacagagg	tgtaccttag	gagctgttc	96420
aaacagtcca tgtggagggt	gtgagtaaga	tgtttccctt	gaatttgc	gaattacttt	96480
tttgggttgg ttgttgg	ttctgagaca	gattctcgat	ctgttgc	ggctggagggg	96540
cagtggcag atcgcgcag	tcactgcac	ctctgcctc	cggttgcag	tgattctct	96600
gcctcagcc cccaaatgc	tggattaca	ggttgcgc	accaaggcca	gctaatttct	96660
tttgtatccc tagtagagat	ggggttcac	catgttggcc	agactggct	cgaactcctg	96720
gcctcgtgat ctgcctgcct	cagcctccaa	aagttctggg	attacaggcg	tgaaccactg	96780
caccgggtcc cttgttaagt	ttattttgtt	ggaagcaaa	ggaggttca	gctttaaaaa	96840
agtttggaaa ttatgtctt	ggttaataat	aaagatttg	gagtaataat	gctttctagc	96900
agaaaagaata aaagaagaaac	agatagcctc	aagaagggg	gccaaagaag	caggctatat	96960
ctgacacact ggggtgtat	aatgggtat	taaaaagaatg	agagcaatga	gcagatagaa	97020
gaggaaattt ggagagtata	ataccatg	gaccaagaaa	gatagactat	caggaaggag	97080
ttgtaaaaat aagttactag	ttctaagaga	gatgttaaga	gggaccggg	aaagccttgc	97140
acaatgtact tagtagcatt	ttacattata	tacatcta	taagaacaaa	tgcgagatc	97200
tcaccattcc tatagactct	tacttgc	tgtctgaaca	cgaaaactgg	ctttgttta	97260
taaataagct aaaattattt	ttgcttcaat	ttctcatgaa	aataaaaata	aaccccttt	97320
taacattgaa aaaatagttt	gaagacagtc	actcttcatt	ttgtatttcc	cacaactatt	97380
attgaatgac tgaatttattt	tttatttgc	agccaaagg	gtgataactg	tatttctca	97440
gactactaaa aatataatttt	atgaattttt	atgtgtctt	atctttttt	gtttttttt	97500
ttgagatgaa gtttca	cgttgc	cggttgc	atgtgtccaa	tctcagctca	97560
ctgcaacccctt cgcctccca	attcaagca	ttctctgc	tcggctccc	aagtagctgg	97620
gattacaggc acctgc	acacccagct	aatttttgc	attttttagt	gagacagggt	97680
ttcaccatgt tggcagggt	ggtcttgc	tcctgac	agggtatcca	cccaccttgg	97740
cctccaaag tactgegat	gcaggcatg	gccaccatgc	ctggcctg	gaatattttt	97800
ctaggttcccc cccaccccaa	gcatttattc	tgcattttt	gttttttcc	taaagcaagc	97860
aagggttaag gattttaaaaa	taatccgtat	tttagaatgc	tttctgg	tgttactttt	97920
tatccacagt agaagttctc	agagaatgat	ctccctctt	taatttaact	tttggcaca	97980
gtattttgag aattataaaat	aatattagaa	tgttttctgg	ctgggtgtgg	tggctcatgc	98040
ctgtatccct ggctacttgg	gaggctgagg	caggagaatc	acttgcacat	gggaggcaga	98100
ggttcgtg atgcgcagg	atgcactgc	atccca	gggtgcacaga	gcaagactct	98160
gtctggaaa aaaaaaaaaa	aaaaaaagag	tgttttctt	cctattttcc	accactgtat	98220
taatgtactt ttccctttaa	gtatttttgc	ctgagtatgc	tgacttaaga	gtaatgttac	98280
aaaatttaat tttaaagttt	ctctgaaagc	cccttatg	gagttttag	ctatcaaatt	98340
gtgtttaattt cttacaattt	tttggaaaaa	ttagtgc	aatatccgt	cattccccac	98400
aaaaaaagcac taaaatcat	gccttgc	aggctgc	accaagtc	gttgcattca	98460
atgcattttc tgccaaatcat	gactctttt	caatgc	gacagccaca	cttagaagc	98520
agccaagcca catggaggcc	gcttatttgc	gtgaccttgg	taagtaacta	tcatttttta	98580
ttaacttgc tttagaaggat	ttgagtttgc	tatgtgaaac	ttctgtcata	ggatacagaa	98640
ctatataattt gggaaatgt	ttggaaaaaa	tgtattttaa	ataacagtc	caagtataat	98700
gggttagtgc gttgtgttcc	ttgttataat	gaatataaa	catggcc	agaaaaacaa	98760
gcattttccat aaaaatata	tctgtat	aaaatataat	atgtaaaaa	atgtctcac	98820
tttattactt agggaaatgt	aaattttaaat	aatcagttaa	tgttcttctt	acacatttgc	98880
atattttta aagtttgaca	atttgcattt	cagtgc	gcagggaaat	accccttctt	98940
tttagtgcata atataatctt	gtgaaatgt	ttggaaagc	aatttggaaa	tcagtataaa	99000
atatgtatgt cattttaggcc	actcttttgc	agacatgc	ctcagatgt	ctcatttata	99060
tgtgcagggtg tttatgttgc	tgtgtgtgt	tgtgtgtgt	tgtatgtat	tgtatgtat	99120
tatgtatgtt tttatgttgc	aggcttata	ttagtgc	gtttgtata	gcaaaaaatt	99180
atggacaaaca tataaataatc	ttgttataat	aaataaccaa	attgtgtat	acgcattgt	99240
tggagtataa tataatgcatt	tttttcttatt	tattttttt	cttgagacag	ggtttactc	99300
tgttgcaggc gctggagtgc	agttgtatgc	tcatggttca	ctgcaggcc	caccccttgg	99360
gcacaagcca ttctctcgcc	tcagcctccaa	gagttactag	gactgcaggc	atgtgtcacc	99420
acaccagat aattttttaa	ttttttgtat	agacagggtc	tcactatgtt	gcctaaatgt	99480

gtctcaaact	cctggcctca	agcaattctc	ccacacaggc	ctcccaaagt	gctgggatta	99540
ccaacgtgaa	ccaccacacc	tggttcatgt	tagccattta	gaaatctaaa	aaagacgtgg	99600
aaaaatgtct	aaggcatgtt	taaatgttag	aaaagcaagt	cacagtatgc	atggtaaaaat	99660
ccgttatatt	aaaataaagt	cttccaaac	aaaaacatat	gcaggagacc	tttattttgt	99720
cagtatttct	tacccaaatt	tctgactta	gaaaattgca	tgtcatgttg	tcataagtgt	99780
aaaaaaaaagat	ccatgaacca	atggacttct	aataaaatca	gtccctgttt	tgacatctct	99840
ctctactttt	gtgtatattc	aaaccagagt	gtcaatgtgt	ttgtggggca	cacttagcaa	99900
taatacatag	cagacaaaat	gcatatagct	cagagagtaa	aattgttaagt	tttgcttagat	99960
cactataaaat	ttgtctatgt	gaattttaaa	ttgtcgatgt	gctctggaaa	acaggcagtt	100020
tctttttttc	tttttttttt	tcttttttag	acagggtctc	actctgtgc	gcaggctgga	100080
gtacagtggc	gtgattacaa	ctcaactgcag	cctcaccctc	ctcagggtca	ggtgatccctc	100140
cctcagtctc	ctgagtagct	gggactatag	gcatgcacca	ccacgcctgg	ctaatttttg	100200
tattnnnnnn	tttttttttt	gtagagacgg	ggtttcgcca	tgttccccag	gctggtctca	100260
aactccttgg	atcaagcgt	ccacttgcgt	aggcctccca	aagtgtggg	attacggcg	100320
tgagactactg	tgctggccct	aggcagttt	tttgggggtt	ttttttgtt	tttattttatt	100380
tgttagacgg	gtctcacagg	ctggagtgca	gtggcccaat	ttttggctca	ctgcaaccc	100440
cgcctcccag	gttcaagcta	ttctccctg	tcagcctct	gagtagctgg	gatgacaggt	100500
gcctggccata	atgcctggc	gatttttgt	tattttag	atatgggggt	tcaccatgtt	100560
ggtcaggctg	gttttgaact	cctgaccatca	ggtgatcagc	ccgcctccgc	ctcccaaagt	100620
gctgggattt	caggcatgag	ccgtcatccc	tggtctgggg	tttcttatga	cgtgaaacat	100680
gcaatttacca	tatgacctag	cagttgcact	ctgtatttt	ccagataaa	tgaaaactta	100740
ccttccataa	aaaacctgt	cacaaatgtt	catagcagct	taatattgg	aaactggatg	100800
ttcttcagca	ggtaatgtaa	ctgggttatt	cataccatgg	aataccattc	agcaataaaa	100860
aggaacaacaa	tgttgataca	tttaaccacc	ttgatgataa	tcaagggaat	tatgtgtca	100920
gacaaaaacc	agtccctaaa	gactacat	atgtatgtt	cgtttggata	atattcttga	100980
aatagagaaaa	ttaagagaaaa	tgaaaagatt	agtgtttgcc	agatgttaga	gacagggagg	101040
tgagaggggt	agtggttgt	agttttaaaa	gtgcaacatg	agggatctt	gtgtatgttga	101100
agttgtatct	ttggcagtgg	tgcagaaatc	tcaatgtgt	aaaatttttt	agaactaaaa	101160
acaagaatga	gtatagatata	aactggggaa	atctgaacaa	tttagaggt	tgatactgt	101220
tcagatctt	agagtatata	tgtaactatag	cttgcaaga	tttttaccatg	ggagaaacta	101280
aagtgtacaa	gggatctct	ggtatttata	tttttttaga	gatgggggtt	cactatgttc	101340
cccaggccgg	tcttgaactc	ctgggtctca	gtgatccgc	tgccccagcc	tcctaaagta	101400
ctggaaattac	aggcgtgagc	gaccatgcct	ggccctttca	gtattgtatc	ttagaacttc	101460
atgtgaatct	agcattatct	catagattt	aatttttttt	aatttttttt	ctcacaagaa	101520
atcagaattt	cctcaagttt	gtgtatgtt	caaagatgaa	ctagttgaca	ctgacagtaa	101580
gactgaggat	gaagacacga	cgtgcttcaa	aaaaatgatt	tgaatatcaa	tggattaaga	101640
agaactctt	tgacaaattt	atgaaaccct	cagtcaatgtt	tataagaatg	cccatcttta	101700
tgatcatgt	atgaaagcc	attttttttt	aaattttttt	tctttcttta	caatttagctt	101760
gtgggttataa	ttaaaatttt	gtttaatata	agataatgt	tttttttttt	tttttttttt	101820
cattttcaaa	gttacgtatc	caaagctact	ctttaaccta	ctatgaatga	ataatgtctga	101880
gttcataaca	tctttgtaga	tatattccaca	atttccctc	aggataatgt	cctacaatgt	101940
gaatttactgg	actgaaaata	atgcagttt	ctaaagacttt	gttatctgtt	cctgaatgtct	102000
cctccaaaaa	gtttttggca	gtttacatcc	tcatgttttt	cgaatgagag	tgttgcctat	102060
tttctgtgtc	cttttttact	gttttaataat	ttttgggggg	aatctaattt	gacagacaaaa	102120
aatgcatttt	atgttaattt	gtttttctgg	gatttttttt	gaggttgagt	atagttttta	102180
atatttttat	ttggcccttt	ggaacttagt	tcataatgtt	tttttttttt	tttttttttt	102240
agtctgggct	gggcgcagtg	gctcacgcct	gcaatcccaag	cactttttgg	ggccggaggt	102300
ggtggttgc	cgaagggtc	gagtttgaga	ccatccctgac	caacatgtgt	aaaccgaatc	102360
tctactaaaa	gtacaaaaaa	tagctcagcg	tgttgggggg	tgcctgtat	cccaagctact	102420
taggaggctg	agtcaagaga	atcgcttga	cccccggaggt	ggaggttggt	tgatttgagc	102480
cgagatcg	ccattgtct	ccagctttag	caacaagagt	gaaaatctc	aaaaaaaaaa	102540
aaaaaaaaaa	aaaaaagaat	ttacatggtc	tgaatttgc	ttaaaagaga	tatggaaatt	102600
attgagtaac	aaataactt	ttaataattt	aggcaatgtt	ttggatgtt	tactttgttt	102660
agaaacccaa	agcatatgt	ttgttagttt	tttattttact	tttagttgt	ggaagtaaac	102720
tttattcaag	gtctctgtt	ccagttgtt	cttttttttt	tttacttata	tttcaatctg	102780
aaatttattt	ttgtctgtt	gctaatttctt	tttcttctat	tttttaggca	tttttttttt	102840
ggactaccag	actcaagaga	ccaaatcaag	cctttcttaag	acccttgaac	aagtcttgca	102900
cgacactatt	gtcctccctt	acttcattca	attcatggaa	cttccggcgaa	tggagcattt	102960
ggtggaaattt	tggtagagg	ctgaaaatgtt	tcatcaaca	acttgggtcg	gaataagagc	103020
acacagtcta	aacacagtga	agcagagctc	acttggcttctc	catctaaaaa		103080

gcatgaaact acagcgtctt	ttttaactga ttctcttgat	aagagattgg aggattctgg	103140
ctcagcacag ttgtttatga	ctcattcaga aggaattgac	ctgaataata gaactaacag	103200
cactcagaat cacttgcgc	tttcccaggaa atgtgacagt	gcccattctc tcctgttga	103260
aatggccaga gcaggaactc	accaagttc catggaaacc	caagaatctt cctctacact	103320
tacagtagcc agtagaaata	gtcccgcctc tccactaaaa	gaattgtcg gaaaactaat	103380
gaaaagttag	tatgtgattt tcttgggtgt	acatatgtgt ctcactttct	103440
tactaaggcg aacttcagat	gaggaataaa atgatggaa	tattttttctt ccctcttaac	103500
tacttgtaaa ttggggagaa	tttggagat gttagagat	cagatcgtg tatggaaaag	103560
gaggcagggt gactggacct	tctaagaagt gtgttatcg	aatttagtaaa tgaagggtca	103620
aatgtcctac ttttccccctc	cactgattt gacatcaaac	cattatccac atagccttat	103680
ttcctccctc ggtttaat	ttattaatat ttactgcac	tttgcagata aaatttttaa	103740
aaaattttta aaaatggcca	ataagtgaca ttatataat	tcagtgccta gtgtatattt	103800
ggattttttt tagtgcac	aagacccctt tgccggact	aggcatgatt atttttttt	103860
ttttgagatg gagtcttgc	ctgtcgcccc ggctggagtg	caatggcgcg gctcggctc	103920
actgcaacct cccgggttcat	gccattctcc tgccctcagcc	tcccaaatacg ctgggactac	103980
aggcgccctgc caccacaccc	ggctaatttt ttgtatTTT	tagtagagac ggggtttcac	104040
catgttcggc aggtatggct	cgatctcccg actttgtgt	ccgcctgcct cgccctccca	104100
aagtgtcgcc attacaggca	tgagccaccc gcccggact	gattatctta ttacacatg	104160
agaaaaaccag ggcttagaaa	ggtagttaa cttccctctag	gttgcatacg aatgtggac	104220
ctagaagcat ttgacaaga	gcacccgttt tttttcttc	tctattagtt tagaaattat	104280
atactcttaa ttatcacctg	ggattttgtat gagacgcct	tcatgttctt ttcatctta	104340
aatgttctt gtgtcttaaa	gggtcaatgtt attttttctcg	atcttttagt tcactcattc	104400
tcagtgact aaaaatggat	ctaactctgc actgaatcaa	gtttcagca ttgtatTTCC	104460
ttcctccctc cttccctctc	tcctccctc aaccaggctc	ccgaggagct ggattacag	104520
gogcccgcca ccactctctg	ctaattttta tatttttagta	gagacggggg ttccaccatgt	104580
ttgtcaggct gatcttgaac	tcctgaccc aagtgaccca	cctgcctcg cccccaaag	104640
tgctgggatt acaggcatga	atcaccacac ctgacggcat	gttattttca tcgcaaaagtt	104700
actgttaagct gggagaatgt	gcacacactt gtaactccca	ctactcaggag acttaaggt	104760
gagaagatg cttgagccca	ggagtttga gaccaacctg	ggcaacacag caagacccca	104820
gctcaacaca agaaaaaaaaa	tttatttttgcatttatttct	atggatcatt ttgttagtt	104880
tcttattctt ttcacccttc	attccactt ttgatccat	cttttattta tttagttta	104940
ttaaatgtat atttgtctga	taatttctgc atctacagtt	ttttgtggac ctgactcagc	105000
atttttttgc ttcttcggat	tcagatgtt gttggcttgc	gatttttagt atttttggcc	105060
gtgaacatgt ttcttgcgt	ttttgtctgtt ggaatttctc	gttactctg tataaattaa	105120
gttacttcag gtgttttgc	ttttcttttgc ctatgcaccc	ggggctctgg tcataccct	105180
tctggtacca cttaaaactg	aatttttgc ttgggtgctc	gtactgatcc tttatgtat	105240
cagggttata cttactgtag	aaatatgggtt ttgattatg	gggtattgtc ccagatggtg	105300
ctggagttt aatatgtct	ctgtttaaaat taatgtgtt	ttccctgtaaa actccaaaat	105360
tctgaattcc agaardactac	ttggccaaaa ttgttaagat	aagggcactg cctgtatTTG	105420
tttctgcctc ccactatttt	ccttagtttta acacaaactc	acctttttaa aaaacatttt	105480
gagagaattc agtattggga	agagtttcta acctgtttct	ggaaatggaa gtccaaagtc	105540
tgtttctgtat ttgtttttt	ttttgagatg gagtctcact	ctgtcaccctt ggtggagtg	105600
caatgacgta ctctcagctc	actgcaaccc ccaacctcccg	ggttcaagcg attctcttgc	105660
ctcagcccc ttgtatgtct	ggattacagg tgcccccac	catgccttgc tgatTTTGT	105720
attttttagaa gagatgggtt	ttcgccatgt tgccaggct	ttgttgcac tcttgcattt	105780
gtgatctgcc cacctcagcc	tcccaaagtgt cttaggattat	gtttctgtaa ttgtataca	105840
tttatttttt ttagaaactg	tttttgcattt agtggtaatt	ttcaataaaaa atagaatata	105900
cagtggagttt attaaaagag	cattagttac atttttccct	tttttattat ctccaaatata	105960
tatataatgt aagtgttgc	ttttttaaaat gtatactgtt	atcgttttta acacatacat	106020
agatttctgt aactgtcacc	actataaggg taaaaggaaac	tttagttctt cacccttggaa	106080
gtcaagcccc acctctatcc	caacacttgg caaccgctga	tctttctccg tctcaatagc	106140
tttgccctttt ctctttttt	tttttatttt tttttttgag	acagcgtctt gctctgtcgc	106200
ccgagctgga gtgcagtgg	gcaatctcggtt ctcactgc	cctccgcctc ctgggttcaa	106260
gcagttctcc tgccttagcc	ttcccttagtag ctgggattat	aggcacgcac caccacaccc	106320
ggctgatTTT ttgtatTTT	tagtagaaat ggggtttcact	catgtttggcc aggtgggtct	106380
caaactttt acctcaatgt	atccacctgc ctccggccctc	caaagtgtcg ggattacagg	106440
cgtgagccac tttttttttt	caggactttt tttttttaaa	tttacattca atttgcatt	106500
ttttttttgtt atggattttt	ccttcagatg cacacctaag	agcccttgc ctaagcaaag	106560
gtcatgaaga ttttctcata	ttttttttttt taaaaggatt	gtgggtggcc aggtgcctat	106620
gettatgcct gtaatctcg	caactttgaga agctgagggt	ggcagattac gaggtcgagg	106680

cggtgcggg gctgaggggc tcctcaacttc gcagacccggg cggctgcggg gcgagggggc	110340
tcctcaacttc tcagacccggg cggccggca gagacgctcc tcaccccca gatgggtgg	110400
cggtcggca gagacactcc tcagttccca gacggggctcg cggccggca gagggctcc	110460
tcccattcca gacggggcg ggggcagag gtggtccccca catctcagac gatgggctgc	110520
cgggcagaga cactcctcac ttcttagacg gnatggcago cgggaagagg tgctcctcac	110580
ttcccaagcg gggccggccgg tcagagggc tcctcaacat ccagacgatg ggccgctagg	110640
cagagacgct ctcacttcc cgacgggggt ggccggccgg cagaggctgc aatctcgca	110700
ctttggggg ccaaggcagg cggctggaa gtggaggtt tagggagctg agatcaccgc	110760
actgcactcc agcctggca acattgagca ttgagtggc gagactccgt ctgcaatcc	110820
ggcaccccg gaggccgagg caggcagatc actcgcggc aggagctgga gaccagcccg	110880
gccaacacag cgaaaccccg tctccaccaa aaaatgcaaa aaccagtca gttggccgc	110940
gtgcgcctgc aatcccgagg actctgcagg ctgaggcagg agatcaggc aggaggtt	111000
cagtgagccg agatggccgc agtacatcc aggtcggct ttcacaactt tggggcattc	111060
agagggagac cggggagagg gagagggaga cgagggagag cccctttttt gcttctttt	111120
ggattatttg aattttccct taaattttt tatcttactt atttatttat ttttttgagt	111180
gattctctg ccacagctcc caagtagctg ggactgcagg catgtgccac tacaccacgc	111240
taattttttt gtatttttag tagagcagg gtttccacat attggccagg ctggcttga	111300
actcttgacc tcaagtgtac cacctgcctc ggcctccca agtgcgggaa ttacaggcgt	111360
gagccaccat gccctgcctt tttctagaat ttatataattt agttcttgc tgatctttt	111420
tatgttaggc ttttagtggc ttctctagga attacaatat acataacttt cacagtgtac	111480
tcacatttaa tattttgtaa cttcaagtgg aatgttagaaa acttaaccac cataaaaata	111540
gaacttaggg tgagggtaaa aaagagagag aaaagaaaatg taataaaatg ttaataaacac	111600
cgtttttttt tttttttctc tttttttttt gagacagatg ctcttttct gttaccaggc	111660
tggagtgcag tggcgtgatc ttggctcaact gcaacctccg cctctgggt tcaagtgttt	111720
ctcctgcctc agcctactga gttagctggg ttacagggtc ggcaccat gcccagctaa	111780
tttttgtatt ttttagtagag acgggttccac tgggttggcc aggtggctc cgatttcttgc	111840
acctttgtat tcgtctccct cagccccc aagtgcgtggg attacaggcg tgaccacccg	111900
cggccggcta agtctttaaa ttttttttgc acatttgcact ttttctttt tccctttagg	111960
attttagtaa cccaaaatgtt agttttgtt ttgtttggca ggttcttgc gcttcttta	112020
cttcttaaa tttttttttc ctgttgcata gtttgcggaaa ttcttattca tctgttctca	112080
aattcaactgg ttcttcccg ttattttccat tctgttattt agtcttgc tgaatttttta	112140
aattttgtttt attatgtttt ttgttctaa aattttctt tttttgtat gtcttataact	112200
ttgtctctga aactttttt tttttccaggat gtatcttcat ttcttagagc atggttttag	112260
tagctactta aattttgtttt tttttttttt tttttttttt tttttttttt tttttttttt	112320
ttgtgagata atgggattttt ctgggttctt atatgacaat taattttggta ttgtatcttgc	112380
gacagtttga ctacgttac atgatttgc atcttgc tttttttttt tttttttttt tttttttttt	112440
agttttgttgc tttaacaagca gttgacctag tttaggttgc tttttttttt tttttttttt	112500
ttctgtccgc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	112560
tctgtgtccca gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	112620
aaggcaggcc atgcacaccc agtcacgcgg tggccccggg agtgcacata caactcgacg	112680
ttttcatggg ctccctctt tttttttttt tttttttttt tttttttttt tttttttttt	112740
cctttatccc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	112800
gcacacttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	112860
tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	112920
gatgtgcctt gggagtcgaa ggagagaaag gaacaaaaca aaacaacca ggggattttcc	112980
tccactctct tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113040
ctgtcttggaa actttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113100
aggaggtgtt ggacaaaactt gttttttttt tttttttttt tttttttttt tttttttttt	113160
tcagttccacc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113220
tgggagagac agaagagtgat gttttttttt tttttttttt tttttttttt tttttttttt	113280
aatcaacggaa tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113340
aaaataacac tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113400
tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113460
ttgtttaaaatc atcaataattt tttttttttt tttttttttt tttttttttt tttttttttt	113520
tttattttcc atttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113580
tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113640
tgagttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113700
ttttccatag tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113760
aaggcaccc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113820
tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	113880

tcttccttg caggtataga acaagatgca gtgaataactt ttaccaaata tataatctcca	113940
gatgtgcta aaccaatacc aattacagaa gcaatgagaa atgacatcat aggttaagcag	114000
tgcttggaaa tatggcaaaa aaaaaatgac aaaaaatgca cagaactgac aattttcggtt	114060
attgactaag ataattttt cttAACATGG aatttagcag ttcccttcct aatttgggg	114120
ctgagtatTTT ttatATCGG attataGTC acTTAAAGG ttTCTCGGT GCAATTCCGGT	114180
cgagggtctt tgccTGGGCC agatGGGTG cAGTGTAGCG ggtgctcagg cctgcccgc	114240
gctgagcAGC CGGGCCGGCG ggcggTACG CTAACCCGGCA cAGACCACG gatggactgg	114300
ccggcagccc CGCACCGTG CACGAAGTGG GCGGGACAGA aACTCTGGG GTTGAAGTC	114360
cagtggaggct AAAAGCCGGT ACCAAAGTCT CTAGGCATCA GGGCTGCAGC CCAAGAGTCT	114420
cacgaccagt gggcaactgg atggccAGAC aggtgtctca gtggtgccct ctccgtctca	114480
gggcttcATC CCACCTCTCA gtgggCCTGA CGTCCTGGG cACCCGGAT gtctacCTG	114540
attagccaga gccatcacat ggccttgac ttgcctttt ttgcccattt attgtggccac	114600
acacagtGTC ATTCTGTG CATTGGAC AGTGGAGGT GCAAGGGAGA GGGCAGCCTC	114660
atgtccAGTC CCAGTTCAc GtaacTTTAT TCTTCTGAAT aaAGACAATT TGCTAACCTT	114720
aaaaaaaaaaa aaaaaaaaaaa agTTTTCTT ATATGTTGGA CCCAAATTCT TAGGCTTTAA	114780
cctgaataac aatgacagca agatcaataa atagTACACA ttTATTAAC ACTCACTGTG	114840
ccccagaca attTCCAAGC ACTTTTATG GATACTGCA ttTTAATTC TAAAGAACCTT	114900
tgtggataa atacagtat ttTATAGATG aAGAAACTGAG AGCAGACAGA AGTTAAGTGC	114960
tttgcTCCAGG GtaacAGCTC AGATATGGCA GAGTCAGGAT TTGAAACTAG ACCTTCACAT	115020
accttaactg CTGTGCTGTG GCAAGTGTGTT TCATACTGTA GGTGGGACc AGCCTCTCT	115080
tatgcCTCA CCCCTGCCA aaaaaaaaaa aaaaaaaaaa AAATATATAT ATATATATAT	115140
atatatATAT ATATATATAT AATATATATA TATATAAAT ATATATATAT ATAAAATATA	115200
tgtattAGTA TATATGCTA TATAGTATAT ATTATATATT AGTATATATA CTAATATATA	115260
atataCATAT TAGTGTGTG ATATATATAT ATACTAGAAAT AAAAATATCA AAGTATCTCA	115320
gagtagtaAG GACAAACATT TCAGAAAAAT GTTTCATTA TATATACATG TATGTATGTG	115380
tatgtgtatt CAACAAATAT ATTCTTATA GGTATAGCA AAATAGTTG AAAGCTTTA	115440
ctgtgtttt TAAGGAAGAC CTTAGTGTAA CGTATATTCA CAGATAAAAG AGGTATTTTA	115500
ttcattcaAT AAATATTACA TTCTCATATA TCCTTAATTT ATGTATTTTT ATTCTTCAA	115560
aaagttAGTA TTGTGATTG ATGAAATAAG ACATGTTCTT GCACTTTAG CAGATCTGTC	115620
ccgatgttgg GCTTCTTAA TCCTTAGTGT GGGTGTCTTG CACTCACTCA CTGCTGGGGA	115680
cagcaagacc CCTGTTAGTC TCAGCTGTGTT CTCATTTAA GGTCTTAAATT GGCCCACGT ACCTCCAGT	115740
tagctattCT GGGGTCCATG TCATGTTGGC TCCATTTCC TTTTCTCTC CCCACACAGA	115800
tacctataAC GGCTATAACA TAGGCCTGGT GGTGTTGGT GGCTTATCCC TATCTGTTG	115860
tattaAGGG GTACTGTTT ACCTGAGTTT GCTGACAGAT GTTGTCTGA GATTGAGGT	115920
tttctgtgtt GTTGTCTAT TTTTATGTGG GAATTGCTA CTATCATCAT CCCTAGACCA	115980
gctttccta GtaatACAAC AGGGATGTT TGACTGATTA GAGTTGCCT GTTGAAGAA	116040
tttggTGGCT AGTGTATTTT TTTTGGGGGG AGTCTGTAC AGTTAATAGC CTGACTGGCG	116100
tgtggataAA AAGGAAGCAG TTCAAGTC AATAAAACAC TAAAATGAA ACCACACTGC	116160
aactctttt CTTTACTTA AGCTTAACCA AATTATGAT GATGTAATCC CATGAAGGAA	116220
aagtcttctg AAGGATCAAG TTGATAACAT TTGTGATCA AAGAATTGA GAAAACCTCT	116280
atcccagtgt CTATCATTAT ATATTTAGG ATGTTAATTA CCTGTGTGGC TTAGGCAAG	116340
tcattttcc TCCCTGAGCC CCATTCTTA TCTGTCCAA ATTATTGTC TCCTCTTGC	116400
gttggactt TTATATATAG CTGTCCTCA AGTGGAGTTT GTTCAAGGA GCCTCCTT	116460
tagctttac TGTGTACCA CTTTGCTAG TCTGTCTTA AATGTAATCC TTGATTTTT	116520
ggtgtgtcta ACTAATTACT GTTTTATGT GAGGATTAG AGTGTACCTG AATTATAACT	116580
tgcaCTACCT CCTTCATCTT CCACAAATGT TTGAAGTGGT AGAATTTTA AAAACTTTGA	116640
aggtacAGCT GACAGAAATT GCTGATGGT TTGAAGTGGAG TGGTATGAGA GGGAAAAAAA	116700
ggaatAAGC ATGACTGCT TTTTGTGTT TTGTGTTGT GTTTTTGAG ACGGAGTC	116760
actctcgca GGTGGAGTG CAGTGGCTG ATCTGGCTC ACGGCAACCT CGCCCTCTG	116820
ggttcaAGCG ATTCCTCTG CTCAGCTCC CAAGTAGCTG GGACTACAGG CGCTGCCAC	116880
cacgcctggc TAATTTTTT TTGTGATT TAGTAGAAAC GGGGTTTAC CGTGTGGCC	116940
aggatgtct CCACTCTCTG ACCTCATGAT CTACTCACCT TGGCCTCCCA AAGTGTCTGAG	117000
gttacaggca TATATATAAG CATATAAAGT GTGTTATAGC ATACAAACAG GTATATATAT	117060
aaacatgcAG TCCACACAGC TGATAGGAAT GAGGCGTAG TGAAGGAGAA GTTGTATGAG	117120
gagaggggac AGTTGTTACA GGAAAGAAGT CTGGAGGAG GAGGAGTGAAGA TTCCAGTGT	117180
cacatAGAAG ATTGCTTAGA TGGGAGCAAG GACAATTAT CTAGACTCAC AGGAAAGAAT	117240
gcagtacacG GGTAGAGATG CAGGTGAGTT GAAAGATGTG AGAGATGATG GAAATAATT	117300
tctgattgtc TCTATATTCT CAAGGAAGCA GGAAGCAAAG TCCTCAGCAA AGAGAATAGA	117360
agaggtgtt AATATTTGAG AAAGGAGATG TACTGTAGAA AAAAATTTCTCAGTTTCT	117420
cTTCTGAAC TCTCACAAAAA CAGAACCTT CCACTGACTCT AGTTGTGTGG GGTTCCTCC	117480

49/122

ctgtcagcta	ccaaattctgc	agatgattgt	tcagtgaaca	ccaactgggt	gtcctctaag	117540
tcagttcagt	tctcacactg	tttacctgga	gatagcatca	gatcccacag	attgaggact	117600
ctgtccccaca	agactgcctc	cacttcagat	gccagtc	agtacaagg	gtggcctgtg	117660
cttctgactg	accccttata	aattggagtt	cccacagtc	cctcccttggg	ttcaataaat	117720
ttgtagtgc	agcttcaga	actcaggaa	atgttttaca	tatattttacc	cattttat	117780
aaaaggatata	acaaaggata	caggatgaa	aggcagatgg	aagagatgca	tgggcaaggt	117840
atggggaggg	ggcacagagc	ttccatgcac	tctccagggt	atgccaccct	ccaagaacct	117900
ctacagattt	agctattcg	aagccccct	ccccattctg	tccttttggg	tttttgtgg	117960
agacttcatt	atataaggcat	gattgtatcat	tggtatggg	tgatcagtc	aacctcagc	118020
ccccctcattc	cgggagggtt	gtgggtaggg	ctgaaagtgc	caaacgtgt	attctgcctt	118080
gggtttttgt	gtgattagcc	tcatcttctaa	agcttttag	aggccacagc	cacaagtcat	118140
ctcattagcc	ttcaaaagaa	tccagagatt	ccatgaattt	taggcgtgt	atgctaagaa	118200
actggctaaa	ggccagttgc	aatgtctcag	gcctgtatc	ccagcactt	gggagggctga	118260
ggcaggagga	tcggttcagg	ccatgagatc	aaaaccagcc	ttgtcaacat	agtgcagacc	118320
ccttacaaaa	aattttaaaa	ttggccaggc	gtatagtc	tgtctgtat	tctcaagtac	118380
tcagaaggt	gaggatcact	gagcccttgg	gttgaaggca	gcagtggacc	atgatcgtgc	118440
cactgactcc	ggcttgggt	aaaactgtgg	accttgttc	agaagaaaaaa	ggaaaaaaaa	118500
aaaactgggc	aaagactaaa	taacatattt	cacagtatca	cagatttgt	ttgtcttagga	118560
aagtgaatgt	aaacagacca	ggacactagt	atgatccctt	gttgcattgt	aggtccact	118620
aaagtcatga	acacaaagtgt	agacttaggc	tcatgttata	tgtttttcc	agccatgttt	118680
aacagctagc	taaatagcta	attgtttcgc	tgcagtttt	tttagcagg	ccttatttttta	118740
gcacatttca	tgtttttaaa	tttcttaccaa	taacattttt	ataaaactttt	ttacagataa	118800
cttcacaaat	ccataatttt	ttaagttaca	atcccagaaa	tagaattgt	cattgaaagg	118860
gtatgttcat	ttttaaagtt	atgctagaaa	ctgccaaattt	gccttcagaa	aaagggtttt	118920
gtatccccac	taaactactgt	gttagttttc	ttgtgcctt	gtcctcaagt	acatattttt	118980
aaaacaatgt	ttggccagt	ttaactagata	aaaggtag	tgccttc	ttcttaactcta	119040
tttgattact	agtgatgt	atgttcttt	cacgttggc	atttttatgt	ttgtccttgg	119100
ttgattgtca	tgcccttgc	tcatttttct	tttggacat	ttcttagtag	tttataagag	119160
ctcttggat	tttaatgata	gtAACCTTT	aactgtcatg	catgtctgc	atcttttttc	119220
tgtttggttt	cctttgtatt	ttgttttttg	agggtttctt	tgtatagggaa	ttaaattttt	119280
tggttggaaa	tcttttgatt	tctgttttgc	catatgtact	tcaaaagact	ttcttattttt	119340
agatcaagt	ttcacatgt	ttttttttgt	ttcttatttt	tttattatggc	tttattatggc	119400
tgtgtgtt	actcccaagt	tgattcaca	gtgtgtat	atagtttggaa	tttagtggca	119460
atthaattat	ttacaacttc	ttttgcagca	aggatttgg	gagaagatgg	acagggtggat	119520
cccaactgtt	tcggttggc	acagtccata	gtcttagtgc	caatggagca	agagaatgtt	119580
agttcatatt	ttcacatgt	gcatacttgc	gaatttgg	tcatgttgc	gaatgggtt	119640
cactcagta	aaaacaaaat	atttttgaga	atthaattat	tttggatatt	tacaagatca	119700
tataaaggat	acttatctt	ggtaacatgt	ttctttttaa	tataaattat	gtgaactctt	119760
aaaattttca	ttttcatttt	caatgttaat	atttcttaag	taaaattat	ttgttttttag	119820
ttctgaataa	atttggggag	tgatttgatc	tgtatgtatt	atgactattt	gaatttggtt	119880
atttattttaa	ataatgtat	tcttcagat	gtcttc	tttgcattgt	aggctttaag	119940
ctaaatggat	gctataataac	taatccaca	tagatttttgc	gaaatggctc	cagaggtttt	120000
tttagattat	tactgtat	tgccctttaaa	aaaaatctat	tcatttttgc	acttaacatt	120060
tatcagaaga	gtgtctgt	taagacgtgg	ttagcatag	tgccagtctt	gaaggaagtt	120120
acagcctaat	aaaagacata	gggcatgtt	tttggttact	gtaatatgaa	gtggcatgtg	120180
ttaaatgtca	ggggagaact	acaaaggat	aaaaagggtt	gagagattac	atacaggtaa	120240
aggaatcagg	aatgacacca	tggggagtt	gttagtgc	acctaggct	ttaaagataca	120300
atagggacag	tatggaaac	gtatattttt	cccactttaaa	cttcttcc	ggtcgttccc	120360
tcaaatttc	cctttgtcc	atgtgcaggt	actttatgt	gtttctgcga	agtcaccatt	120420
tctgtaaata	ccagattgaa	gtgctgacca	gtgaaactgt	ttacctggct	gacattctct	120480
tctgtgagtc	agccctctt	tatttctctg	aggtaaaagt	tgcatttctt	ttcacactct	120540
attcgagcat	tccagcttc	aactatcaat	gctggggccc	tgtctatagg	aaataaacaca	120600
gaagaggccaa	gtcatttcca	aaaagatgt	tcatttttgc	aaatgttttc	tgatggcaag	120660
agtaattttaa	taatattata	gagagaacat	gaaaatttca	tgttattttat	aactcttaatt	120720
ttgagaaacc	taattttact	actgcattgt	agagagtgc	tgtttttat	tatggagc	120780
tatttttaaa	ccacagaatt	tgaaacttgc	ttccagtgca	taaatttgcag	accagacttc	120840
agaagagaaa	aaaagtgt	aattttttct	tatgtctatc	atttttactt	tagtacttgc	120900
ataggatgc	ccagtgaaga	agcatgttgc	acagacatgt	agtatattaa	tctttttggag	120960
gcatacagt	tagtataat	ctctttgtt	ggcttcaaca	agtgtaaatta	ttttgttggaa	121020
aagcaatgt	ctattaaatgt	gaaagaggat	tcccagtc	acaaaggcgt	aatttagaca	121080

50/122

ctcgattctg cctctttaca agaatacagg tactcagttg atttgtttc tcactccctt 121140
 tctttgcstat aagttaaat caacaattt gtttaggttaa tatgtcctca tggaatggtg 121200
 gaaatgatca gatataaaaat atttgggttgc tttagttac tctttatatg tttgctggca 121260
 aggaaccaca aatccagttt agtataattt ttactcttagt tcactaaaag tttgcattcca 121320
 gctgtgttagg tagtgtttgt ttcttgcattaa ctttttttc gtctaaaaga atactttaaa 121380
 actttcaat ctcaaatttgc tgtaacttgc tgacagggtgt taacagaaga agtagatctt 121440
 tttgtttttt gcttatgacc tggattttaa tattttgcatt tatagatgg agattgttag 121500
 agaaaatctgt ttatagttctt attttccctt gtgtttttt tcttccttagt acatggaaaa 121560
 agaggatgcg gtgaatattct tacaattctg gtggcagca gataacttcc agtctcagct 121620
 tgctgccaaa aaggggccaaat atgatggaca ggaggcacag aatgatgcca tgattttata 121680
 tgacaaggta gttatattga tagatggatt cagcagatac ttattgaaca ttgatatatgt 121740
 ttttgtggaaa taaaatgaa taaaactcagt ctctgttgc aaggagctca caggaggcag 121800
 cataaaaagct gcttttatattt ggtgtttgt aagctttttt ggttctttaga aaaaaaggttt 121860
 ctgctgtggaa aggggagggtg tatgtgggtt aaacaggatg gcaatggttg tttcaaggaa 121920
 gtgtttccca gaagagagat ttgtttggaa tcccaaaagaa agaaggaaat ttgcttaccc 121980
 agagaaggca gaaaacaaca ttctaggca aggcatggc ccagaagcca tgaaacgta 122040
 ggggaaaggta gcactttcaaa gaaaacttgcg ttttagataat caaaggagtgc gggaaataaaat 122100
 attaggatgc tggtaatattt tggatagat tggtaaggac cttgaatgcc tattttatggg 122160
 tatattatac ttctgttata aatctgc tggcacgttgc taatagttt ttatattatgt 122220
 ttcaactgaaa atgagaggat gggaaacatca tacagtaaac aaaatttggaa atatctggc 122280
 aggcagatga tgagcttgc gcccaggctg taacgtatgg tatttttcc attaacttt 122340
 tcttactctg taaaaaaaat aattcgttgc cgggcacggc ggctcaacttgc tgtaatcaca 122400
 acacttttgcg aggccaggac aggttgcattt ctggagcccc ggaattttgc accaggctgg 122460
 gcaacatggc aaaaacccgc ttactaaaa atacaaaaat tagctgagcg tgatggcgtg 122520
 cgcctgttgc cttagctact tagggccctg aggccaggatc atcacccatc ctttgggagg 122580
 tcgaggctgc agtggactgt gatccactgt actccacccct gggcaggcga gtagagttag 122640
 accctgtctc caaaaaaaaaaaa aaaaacaaca aaggttataat ttatttttgc ttcttaagca 122700
 aatgtctaaag gggtaactttt gggatagaa aagttccaca gatgtttggg tttaagaca 122760
 ctaatagttt cttagggcattt ggttcttgcattttgc tgggtcttgc tgggtctgtc 122820
 tgcataggaa tcaacctgaga gcttattttt aataggtttt caggcttgc ggggtggctc 122880
 acgcctataa tccccagcaact ttgggaggct gaggccaggcg gattacttgc ggtcaggcgt 122940
 tcaagaccagg cctggccaaat atggtaaaaac cccgtctcta ctaaaaatac aagaatttagc 123000
 caggcatgtt ggcacacacc tggtaatccca getactcagg aggttgcaggc agggaaatttgc 123060
 ctgcgagcccg ggagggtggag ttgcattgtc gcccaggatc tgccacttgc ctccaggctg 123120
 gctgacagag ggagactctg tctcaggaaaaaaa ataggtttt caggcttgc ggggtggctc 123180
 ccgggtggctc acacccatgttca tccccagcaact ttgggaggcc aaggccaggca gatcacttgc 123240
 ggtcaggagt ttgagaacttgc cctggccaaat atggtaaaaac ctgtctcta ctggacttgc 123300
 caaaaaaaaaatc actggggattt ttgacgggtt cttataatcc cagctacttgc ggaggcttgc 123360
 gcaggagaat tgcattgtc gggaggccaggc agactgtatc ctcaaaaaaaa aaaaaaaaaaaa 123420
 aaagggttcc agtccccctgt tctcaggaaaaat tctgattttgc cagggttgc gttgttgc 123480
 gaatctttat tttagaaaga cataccagat aattctgtata aatagccatc tttagggatgt 123540
 agtcttaattt tccttattttt caagtttgc aaaaatggcc cagagatgttgc atgatatttttgc 123600
 caaaatgcaca gaacaaggta gttggccatc ttggacttgc atgcatttttgc taatgttctg 123660
 tccatgtttt attctgttgc atgttttttgc tggatgttgc atgcatttttgc taatgttctg 123720
 atagatgttgc agataggaaat agtttacatt tggaaattttgc tggatgttgc atgcatttttgc 123780
 caagtcgtgg aggaggatattt accaacttac tcaatatacaac ataggagatt cacattttgt 123840
 tacaatatttgc ctgatatttttcaaggatgtt tctttttttt cttctttttt attttttgc 123900
 atggagttttt gctctgttgc ccaggatgttgc tgccacttgc acgttgcatttgc ctcaacttgc 123960
 cctccacccctt cttgggttcaaa ggggttcttgc tggctcatttgc tggatgttgc tgggttgc 124020
 aggtggggcc caccacccgc agttaatttttgc tggatgttgc tggatgttgc tgggttgc 124080
 tggatgttgc ggggttgc tggatgttgc aactccgttgc ctcaaggatgttgc ccaccacca ctggctccca 124140
 aagtgttgcatttattataggcg tgaggccatc tggccaggatgttgc gtttgcatttgc tggatgttgc 124200
 tatgcattcat cataatcatc cattatcaac ctttgcatttgc tggatgttgc tggatgttgc 124260
 tttagatgttgc tggatgttgc ctttgcatttgc tggatgttgc tggatgttgc tggatgttgc 124320
 catttcataa tcaatatttgc aaaaatatttgc tggatgttgc tggatgttgc tggatgttgc 124380
 ttgtgttatttgc aatatttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc 124440
 tccaaatagg tcaatatttgc tcaatatttgc tggatgttgc tggatgttgc tggatgttgc 124500
 atgtttacttc ctgatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc 124560
 aagtgttgcatttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc 124620
 ctatctatca agtccctcaac agagaatagg taccatataa tggatgttgc tggatgttgc 124680

tgcctcagtt	gtagtctgat	ccttacagct	tttaaacaac	agtagagttc	accgtcaaga	124740
actaaggatg	gttggcaggc	agatagaaaag	gtagcaagtt	gacccaacta	tctctggga	124800
agtggaaaca	aagaaagggtt	acatcagcac	tgcatacaca	tagctctata	gttctaggcc	124860
tgcaggctca	atcaagtagc	cttgataaag	attctctgga	ggaggtgctg	aaagttgctt	124920
atacttgcata	ttgaaatttga	tttacttcg	gatatcttt	taccataggt	acttctccct	124980
ccaagccaca	catcctcttg	gatttgcata	tgttgcata	ttagaaattt	aatccaatat	125040
ctgcaggaa	ggtggggccac	tccccaactt	tttcacaaact	ccattacgtc	aggcctggac	125100
aaccatggag	aaggttaaccc	agaatctca	acgtatcaa	ctacaagaag	tttatttggt	125160
agaactcata	aaataataagg	ttggaaaacc	aagcagaata	gcacagtgg	aattgaagca	125220
gtccggcaaa	gtgatttgcata	gcagaggct	tgagtctggc	ctggtatgt	cagtacgtg	125280
ccacataaca	tttttagtcaa	cagtgcact	cgtgtacgt	ggtcctgtac	gattataatg	125340
gatcaaagct	ggtagtgcaa	taataacaaa	agttagaaaa	aataaatttt	aataaagttaa	125400
aaagaaaaaa	gaaaaaactaa	aaagataaaa	gaataaccaa	gaacaaaaca	aaaaaaattt	125460
taatggagct	gaaaaatctc	tgttgcctca	tattttactgt	actatactt	taatcattat	125520
tttagagtgc	tccttctact	tactaagaaa	acagttact	gtaaaacagc	tccagacagg	125580
tccttcagga	ggtttccaga	aggaggcatt	gttataacaa	gagatgacgg	ctccatgcgt	125640
gttactgccc	ctgaagacct	tccagtgga	caagatgtgg	aggtgaaaaga	aagtgttatt	125700
gatgatcctg	accctgtgt	ggcttaggt	aatgtgggtg	tttgcctttag	ttttaaacaa	125760
acaatattaa	aaagaaaaaaa	aaaataaaaa	atagaaaaaa	gcttataaaa	taaggatata	125820
atgaaaaat	ttttgtacag	ctgtatatgt	tttgttttta	agctgttat	acaacagagt	125880
caaaaagcta	aaaaaaagtaa	aacagttaaa	agttacagt	aagctaattt	attattaaag	125940
aaaaaaattt	taaataaaatt	tagttagcc	taagtgtaca	gtgtaaatct	acagtagtgt	126000
acaataatgt	gctaggcctt	cacattact	tacactcac	tcgctgactc	accagagca	126060
acttccagtc	ttgcaagctc	cattcatgtt	aagtgcctta	tacagatgt	ccattttta	126120
tctttatatac	ttgttttta	ctgtgcctt	tttgttattt	ttgtttaaata	cacaatttct	126180
taccattgca	atagtgccct	acgatattca	ttatagtaac	atgtgataca	ggtttgtac	126240
ccaaaagcaa	tagttgtac	catatagcc	agggggtgtag	taggcatac	catttaggtt	126300
tgtataagta	caactctgt	tggttgcaca	atggcaagca	gcctaacgga	aattctgttt	126360
attgattgtat	ttgttgcata	ttgttgcata	acagatgttc	actccattgt	ccaggctgg	126420
gtgcaggatgc	acagtctttg	cacactgca	cttgcgcctc	ccaggttcaa	ccattatcc	126480
tgcctcatcc	tcccaagtag	ctgggattac	aggcaggcac	caccatacct	ggctaatttt	126540
tgtatattttag	tagagacagg	gtttcaccat	tttggccagg	ctgttctcg	actcctgacc	126600
ttaagtgatc	tgcctgtgt	ggcctccgg	agtgtgggtt	ttacaggcat	gagctaccat	126660
gcctggggcag	taactgaat	tctctaatgc	cattttcctt	atctgtttaa	tgacgataat	126720
atgcacgtt	acctcaaaat	tactttgtat	attaaatgtaa	ggtaatgtat	ataaaatata	126780
tattaacata	gtacctgaca	catggtaagc	atcaaaaaat	gttaactact	tttattacta	126840
ttattattac	gtatttttaa	ataatttagag	agcagttatca	aaaatttagt	gggcgttagt	126900
gcatgcaccc	atagttccag	ctactccgg	ggctgttgc	ggggaggatgc	atgacgctgg	126960
gaattaaagg	ctgcgttgc	ccgttgcata	ccccctgcac	tccagcttg	gtgacagagc	127020
aagacctgt	cttgaaacaa	taaaaaggc	attatgcgc	aacgttagt	tagaaatgtat	127080
ccacatataat	caccagtaac	tgtcaacagg	atttggaaacc	tagtttggg	tattatgatc	127140
acaaggatt	attaatagct	tattaataat	aaagcggttgc	ctaggcacgg	cgactcacat	127200
ctgtatccc	agcacttttg	gaggccgagg	tgggtggatc	acctgttgc	aggagtttgc	127260
gaccaggctg	accaacatgg	agaaaaacccca	tctctactaa	aaatacaaaa	ttagccgggc	127320
gtgggtgtgc	atgcctgtaa	tcccaact	tttaggaggct	gaggcaggaa	aattctttgt	127380
acccgggagg	cagagggtgc	agtgtactgc	gatcgaccca	ttgcactcca	gcctgggca	127440
caagagcaaa	actccgttcc	aaaaatataa	ttataataaa	taaataaaag	taaagtatttgc	127500
atgtttgtat	atgattttt	tttctataat	actagaggag	atttttccgc	gaattttcaga	127560
gccagtgggg	ttatgttgc	tgtatgtgc	atgtgtatcc	aggtgaaaaa	acttaattaa	127620
acgttattat	ataatccat	acataaaaac	tgaatttttg	gaatactgaa	gaatgacata	127680
tagaagtcaa	atcattaaat	agcttagtgc	aaacagaata	gagtgtcgc	tgttacccaa	127740
tgatgataat	attttcacga	ttaaaattaa	accttttctg	attttaaagg	aaaagttcag	127800
atctgtatca	tataaagaat	gtaaaatttc	agggtaataa	aattttaaat	cagagagaaa	127860
aatgcaaaaaa	tagttcttac	tagatgtgt	tatgttaagga	acttagacta	attttaagaa	127920
cactgtcaag	accctgtgt	tttaggttgc	aaaaagacat	gaatgattca	ttcaacaaaa	127980
actttgatgt	tttctgtgt	agatgttgc	tttacagtgg	taaacaat	aaatgtgttt	128040
ctgtatccct	ggagcttagt	ctacaaaaaa	ggtacatatt	ggccgggcac	ggtggctcac	128100
gcctgtatcc	ctagcactt	ggaagatcga	ggcgggggtgg	tcacctgg	tcaggagttc	128160
aagaccagct	tggccaaacat	ggcgaacaccc	cgtctctact	aaaaatacaa	aaattaactg	128220
ggtgtgggtgg	cgagacaccc	taatccacgc	tactcgggag	gctgaggcag	gagaatcact	128280

tgaacctggg	agacagagg	tccagt	cgagatcatg	ccactgcatt	ccagccccggg	128340
ggacaaaagc	gaaaatacgt	ctcaaaaaaa	caaaaacaaa	caacaaggc	acgtattaaa	128400
tacgaacata	aatatttaca	aattatactg	aataagttct	catgtttatt	atttgcttgc	128460
ccagttacaa	acttttcctt	cgtagaatta	gaaatataaa	taataaacat	gagaactcat	128520
tcagttataat	taataattat	taaatgtaaa	taaaaacatc	tatgtacaat	taggcattta	128580
tttaagaatt	atttgaaaaaa	aaaacaatgt	ggaaacagat	atttgtat	atttgcttagt	128640
attgaaaattg	ataatgttt	tttgaagagt	aaagtgacca	tatataat	agttaaaatt	128700
taactcagca	atcacacgccc	ttgttggat	tottaaggaa	atcgtttga	aagttaaaatc	128760
aatatgcata	caaagacttt	aacattatc	ataaaaccaga	aaaatcgagt	ttccaaattat	128820
atcctatgg	ctatttctg	ctaaaaagta	ttaatatcaa	ctttatgtaa	tactttcg	128880
acaatattt	ttgggggagaa	aacccaacaa	aattacatgc	attgtat	ttttttttt	128940
tttttttta	gacagtctt	ctccagcgtc	caggctggag	tgcagtgt	caatctcg	129000
tcactgcaac	ctccatcttc	caggttcaag	caattcttc	gcctcaggcc	tcccagtag	129060
ctgggattatc	aggcgctc	caccatgc	actaatttt	tatagttt	agtagagatg	129120
gggttcatc	atgttggco	ggctggctt	gaactcctgg	tctcaagtga	tccgtctg	129180
tcggcctc	agagtgt	gattacaggt	gtaaggccact	gcacccagcc	ttatgcatta	129240
taattttat	ttgttaactg	tacaaaggga	taataacttgc	agtacacaa	gaagttaaaaa	129300
catttttat	aggtgttta	catttgcac	tagtgcatt	ataggtaaa	tttattttt	129360
taaaacatgt	ttatgtggat	ttgatttca	ctttaaaaata	atgtttca	tctctatcag	129420
gtcttttgc	ctggctttt	gtccagcaat	cttatttata	aatattgaa	tgtatctcatc	129480
cattcggtt	gaggagatg	atttctggc	gggaacgtgt	cgctgactgc	tcctggctct	129540
gttggccctc	ctgatgagtc	tcacccagg	agttctgaca	gctctcg	tcaggat	129600
actgattcgc	tctgcattt	gggagaaag	catacata	tttcccttca	catcccagta	129660
acagatctt	ttatgtgtt	attttgc	gtggaaaaaa	aagataaaag	ccaggcacag	129720
tggcctgtgc	ctgtatccc	agcacttgg	gaggctgcgg	tggcgatc	acacgagg	129780
aggaattcga	gaccagcctg	gccgacatgg	tgaacccccca	tctctactaa	aaataaaaaa	129840
attagccggg	catggtggc	ggcacctgt	atcctagct	cttgggaggc	tgaggcagga	129900
gaatcgctt	aacccaggag	gcagagg	caatgcacca	aatcagcc	actgcactcc	129960
agcctgggt	acaaaggtag	actgttctc	aaaaaaaaaa	aaaaagaga	gaaataaaat	130020
tagcttactt	actatctt	aatcaagca	tttgtggta	cttaaaat	actgtattt	130080
aaagtatcat	gctgttcat	ttagggcatt	attctat	aatctgtggc	tgtttctt	130140
aataaatcaa	gtaatatg	atataatcat	agcctctgaa	gagcttta	tgttaagtatt	130200
tatttagat	acttttttgc	aaataatgt	atgaatttct	aggctctt	tttttttctt	130260
ttcttgagac	agggtctc	cgctgc	tggaaatttct	gggctcaat	aatccaccca	130320
ccacagcctc	ctgatagat	gggacttagag	gcatgcacca	ccacgc	ctaatttga	130380
atttttttt	ggccaggcat	gatggtc	gcctgtatc	ccagcactt	gggagaccga	130440
ggcaggcaga	tcacgagg	gggagatgg	gaccagc	gcacatgg	tgaaccccc	130500
tctctactt	aaataaaaaa	attatgt	tatgttgc	catgc	atcccagta	130560
cttggagac	tgaggcagg	gaatggctt	aaccaggag	tggaggttgc	cagtgc	130620
agatcacg	actgcactt	tgcatgt	cagatgaga	ctccatctca	aaaaaaattt	130680
ttttttttt	tgatggagtc	ttgtgtt	gctcagg	gtcttga	cctgac	130740
aatggccct	gtttcagct	aaagt	tttttgc	aagagacagg	gtcttgc	130800
gttggccagg	gtatgttca	actcttgc	tcaagc	ctccac	ggctctca	130860
agtgttgg	ttacagg	gaaccact	ctataatgt	gtt	caaggc	130920
tgatttgc	ttgcattt	gtgcacatt	gtcattt	ttgac	ttgggtttt	130980
tggagtgc	tcatatgtt	aaccat	gattctt	aaaatc	aaatcaca	131040
atcctaagac	aagaaatcta	aggagg	aaaatgtt	ctgg	ttaactcaca	131100
cagttaaatgt	tagaggcaga	aatatccc	ttctgtt	cttcc	agttatgt	131160
agcataataa	tttcttatt	actgttgc	aaaaatataa	ccttgc	ttaataact	131220
ggccttggat	aaatttctt	aatttgt	agatgtt	ctgttgc	tttacaaagc	131280
tctctgt	atctttttt	tctgtt	ttacctt	gataaaact	tgatgtt	131340
tgccat	tttttttgc	ataggtt	atgtt	ttttttt	ttttttttt	131400
gttttgc	ttgttgc	ttgagac	atcttgc	gtcgc	ctggag	131460
atggcacat	cgtggctc	tgcac	acctcccgg	tcc	acttgc	131520
cagttctt	atgttgc	attacagg	tgtca	accat	tctctgc	131580
tttagt	gaaggggtt	caccatgtt	gtcagg	tatt	ctgac	131640
gatccac	cctccgg	ccaaatgt	gggat	gc	ctgcac	131700
ctgaatgt	tgttttgc	taggcactt	agaaaggc	atgtacta	acataaaat	131760
attttatac	cttttgc	tactatata	atagaaaact	gcactt	atcattt	131820
cac	aatgttca	agcagaacta	cccatgt	cccag	atgc	131880

53/122

cagcattatc	agccccctcta	gaagccctct	tgggccccctt	ccattcaactg	tccttcttgc	131940
caccaggta	gtactatcc	tgacttttga	tgcatacat	tagcattacc	tgttcttgc	132000
attttataaa	taaaaccata	ctgtgttattc	ttttcttgc	cagcttattt	gtgtaattc	132060
acatttacat	catacaattc	agtgggtttt	atatggtcac	agagtttagt	aaccattacc	132120
acatcgattt	tagaacattt	ttttcaactcc	agatagaaaac	ccccttactt	taaactccaa	132180
atcccccaact	ccaccagccc	taggcagcca	ctagtcaact	ttttatcttct	ataagagacaa	132240
tagatttgc	tattctggac	atttcaaaa	catggaaaccg	tatattatgt	ggctttttgt	132300
tgc当地	cttcaactta	gcatcatgt	ttcaaaaagag	catcatgtt	tccatgttgc	132360
gcatgtatca	gaattttattt	cctcattatg	gccaaatatc	ccattgcaag	gatttatgac	132420
attttatttgc	aattgtaccc	tccttctgc	catttatcaa	taatgtcaact	gtgaccattt	132480
gtgtacaagt	ttttgtgtgg	atacagggtt	tctttttgtt	ttttaatttgc	aggtggagtc	132540
ttgtctgtc	ggccagggtc	gagtgcgt	gcacaatctc	ggctcaactgc	aacctctgtc	132600
tcctgggttc	aaggcgttct	cctgcctcag	cctcccgagt	atctggact	ataggcacgc	132660
accaccacgc	ccagctaatt	tttttagtgc	gatgggggtt	caccatgttgc	gccagcttgg	132720
tctcaactc	ttgaccccaa	gtgatccacc	catctcgcc	tcccaaagtgc	ctgggattac	132780
aggggtgagc	cactatgccc	ggctgtgggt	ttcattttctt	ttgttgcata	tacataggag	132840
tagatttgc	gagtcaagag	gttaacttca	aacttattga	aaaactgc	gatttttttc	132900
cggaaaaggct	gcaccatttt	gcaatccac	cagcgtgt	tgagttttac	agcttctcca	132960
catttcatttgc	gaacttatta	tctgttttgc	ttttttttaaa	aatgatagtc	atccaataaa	133020
gttctacttc	agtgtgggtt	ttgcacttct	ctgtgtgc	atgatgttgc	gcatcttttc	133080
atttgtttat	ttgcctttgt	tctagtttgc	aaaaatgtt	tatcaatac	cttggccat	133140
tttttattttt	atttttttttt	attttttttt	tttgagacc	aagtctca	ctgtcagcc	133200
ggctggagta	aatatgtgt	gtctcagtc	actgcacact	ccgcctctg	tgtcaagtgc	133260
attctcctgc	ctcagccctc	cgagtagct	ggattacatt	tcaggcacct	gccagcatgc	133320
cgggctgatt	tttgttatttt	tactagtgc	agggtttc	catgttagcc	aggctggtca	133380
caaactcctg	acctcagggt	atctgcctgc	ctaggcttcc	caaagtgc	ggattacagg	133440
cgtgagccat	ttggcccccagc	ctagattttt	tttttttttgc	tttttttgc	gaaggagtc	133500
tgctcttgc	ggccagggtc	gagtgtat	gcaacatctt	ggctcaactgc	aacctctgc	133560
tcctgggttc	aaggcatttt	cctgcctcag	cctcccccagt	agctggatt	acaggtgcct	133620
accaccacac	ccagctaact	tttgcattttt	ttttagagac	agggtttc	catgttgcc	133680
aggctggct	caactcctga	cctcagggt	ttcacctgc	ttggccccc	gaagtgc	133740
gattaccggc	atgagacttcc	aggccaggcc	aaatttctca	ttatattgc	caggctgg	133800
tcaaactct	gggttcaagt	gatccctctg	ccttggc	ccaaagtgc	gggagtagac	133860
gcgtgagcc	ccttgc	cccccttgc	cattttttaaa	ttagattgc	tttttatatt	133920
gagtttcagg	agtcccttta	atatttgc	taatgtccc	ttatcaatt	atattatttc	133980
caggatattt	tttcattctg	tgagtgttgc	tttctctacc	ttttttttttt	gggggtttttt	134040
tgttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	134100
gcagtggcac	atcacagct	cactgcacc	toaacttctt	ggggcaag	gatcccttca	134160
cttcagccctc	ctgatagact	agggccatag	atacacacta	tcacacccag	ctttttttttt	134220
ctgtttgttag	agacagatct	tactgtgttgc	cccaagttgg	tctcaactc	taggctcaaa	134280
gtgattctcc	cacctctgc	tcccagagt	ctgggattac	agggtgc	cacacgc	134340
ctgttcttcc	actattata	gtgttcttcc	gttccgc	cccgagtagc	tggattaca	134400
ggcacccacc	accatgcctg	gtaattttt	ttgcattttt	tttgcattttt	tttgcattttt	134460
atgttccccc	ggctggtctt	gaacttctga	cctcagggt	ttcacctgc	atggccccc	134520
aaagtgttgg	gattacaggg	gtgagccact	gacccggcc	aaaatattgc	tttcttaaca	134580
gtattgtctt	ctaatttgc	aacatggat	tatcttcat	tatttatgt	tttcttctt	134640
tcagcagaat	tttgcattttt	tcagactaga	aggcttcc	cttcttgg	cattttattcc	134700
tattttttaa	tttcttttgc	atccattat	aatagaatt	ttttttttttt	tttcttctt	134760
agatttttg	atgagagagc	atagaatatac	aagtgttgc	tatcatgt	tcttgcact	134820
tcaacttgc	taaatctgt	tgttagctt	aatagttttc	ttgtggattc	tttaggattt	134880
tcaatataata	agatcatgtc	atttatggat	agagatagt	ttttttctgg	ctagaactt	134940
cagagcaatg	atgagtagaa	gtggcagaag	aaaaatctt	tgtcttgc	cctatctgc	135000
aggggaaagct	ttcagtttca	tcatttata	tgatgttgc	tgtgggttt	caataaatgc	135060
cttttttgc	attcaggat	ttccctatca	ttctgttgc	tttaaggctt	tttttttttt	135120
ttaaatcatg	aaagggttt	gaatattgtc	atgttcttgc	tgtatcagta	taaatgatcc	135180
tatggatattt	gggtttttt	ctgttgcgt	gaaaatattaa	ttgatttca	gatgttaaac	135240
caacccctgc	tacctgagat	gaatctca	tgttgc	gtataatctt	tcaaatatgc	135300
tgctggatcc	catttactgg	tatgttgc	aagatgttgc	atctgaa	ttaagataac	135360
atttacactc	tatcagaata	gaatttgc	taaatgtg	agtgat	ttgggttcttgc	135420
attcttcc	attccaaaga	tagacataca	tccgtctgt	tgtctgtt	tatgccc	135480

ccatactctc ttgattacta ttgcttgt aataagttt gaaatcagaaa gtataaatga 135540
gattttggta tctgagtaac agtcctcata gaatttagttt gggaaatattt cctctttatt 135600
ctggccctc tttttttttt gtttaactgt gtatcttggaa gattgttcct tctcaacaca 135660
tgagagccgc ttccccatcc ctccccacccc tgctatagag aggtctataa gtgtctgtc 135720
aattatttta tttacttaac ctattactta gtcggggaca ttaagcttgt ttatgtctt 135780
tatTTTAAAC aatgtcgac tgaataatct tgatataag tcattttcca tcaatataag 135840
tctctctgtt actgaatttt tagaagtggaa atttcttagt caacctatgg ctctgtatTTT 135900
cacaaaaata ccaattctgg tttttttttt ggagggtgggg agtagggatg agaatgtcg 135960
aggagaaccc gctgtacta gctggctgt cattttgaa aggtttctt agcttcttt 136020
tgtcatatgg cctcaccaag aatcaaaaac attccttattt accctgtaaa catggggctt 136080
tactacccaa gatacatatt tctggatgt tgacagctt tcattttgaa gaaataatgc 136140
tgtgagtaca gcacattttg tggaaacttag gtcgtttaaga atgtcttata aatttcataca 136200
ttatCACATT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT 136260
aggccagcgt gcaatgggtac aatctggct cactgcgacc tccatcttcc gggctcaagt 136320
gatttcatg ttcacccctc cagagactt atggttacag gcatgcacca ccatgcccc 136380
ctaattttttt tatttttagt agaaactggg tttcaccata ttgaccatgc tggcctcgaa 136440
ctcttggcct caagtgtatcg gcctgcctca gcctcccaaa gtgctggat ccttgtattt 136500
ggtaaaaagat gaatatttgg ggctgtatcg tggtctatc ctgtatccc agactttct 136560
gagactgagg tggggaggat cctggagccc aggagggtga ggctgcagtg agttgtgatc 136620
gcgcatttgc acattcaacctt agaaattata ggcttcagtc actgtgccc gcatgtacat 136680
ttaatattt tgccttcctc ttttagctat agtatgaggt tacatttcag agtattttttt 136740
gttaagcatc ttaatagtgt tgaggttgag tgaaatgtt ttcttatttca aacactgaag 136800
aaaattttgtt acaaattttgtt cacattccaa gcccaggact gattgttca tataatttcta 136860
atTTTACAAT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT 136920
ttttagtgaag cgataatttttggatggc aagtctggatc cagaatctt atatcaacagg 136980
acatatgcgg ggttaagctt gctcatgcct agaatttttca caagtgtaaa taacttttgc 137040
tcttttaaat ttttttaattt aatttttacat ttttttttttca tcttatttttta tatgcccaga 137100
acttttcaactt agatgtgtca gtataatgtt tggttttttca aataaggctc tggagtgtact 137160
tccttgggtttt taattttttttgc tctggcttatttggcagcc gctaaccctt tggtatctca 137220
gttttttcat ctgttttttttca agaataatataa agtggaaatggaa tgccaaatc atttactctg 137280
ggctgcataa ctgttttttttgc gaaaaaggat tcccttttttca ttaagaatttta agttggat 137340
tcattttttagc ttgtatTTTTT aagatgtgtca ttcataggat atgcccactt ctgaaattttt 137400
ccacagatcc aatcttttttttca tcaatttttttca ttttttttttca atagctgtat taacatgtaa 137460
agggtgtttaa aggcttttttttgc ttttttttttgc gatccgtacc ccgttttttttca gcagcaccat 137520
gaaaaaggaa ttcaacatca tttaatttttttca ttcaatgttttttca agtaacatgtt tgaggccagg 137580
cgcagttggct cacacccgtt aatccacccac ttttttttttgc cgagggtggc ggtacacgag 137640
gtcaggagat tgagaccatc ctggcttacca cgtatggggcc cccgttcttac taaaatataca 137700
aaaaatttttttgc cggggatgtt ggcaggccat ttttttttttgc gtttttttttgc ggttttttttgc 137760
aggagaatgg cgtgaggccgg gggggggggat ctttgcgttgc gtttttttttgc tgccactgca 137820
ctccagccctg cgcagttggat cggagacttttgc ttttttttttgc gtttttttttgc gtttttttttgc 137880
tttggggagact gaggagccca gaaatgttttttgc gtttttttttgc gtttttttttgc ttgaccacat 137940
tgtatTTTTG AATGGGGAGCTT TGTATTTTTT TGTATTTTTT TGTATTTTTT TGTATTTTTT TGTATTTTTT 138000
cttttagaaaa ataacggaca agggccggca cgggtggctca cggctgtttaat cccaccactt 138060
tggggggccg agacccgggg attacccgtt ctcaggatgtt ccggacccatc ctggggcaaca 138120
tggtaaaacc cccgttcttac taaaatataca aatgttttttgc gggcgtggat ggtgtgcct 138180
gtagtcccaat ctactccggat ggcttggatca ggagaatttttgc ttgaaaccgg gggccgggg 138240
tttgcagtgtt ccaagatcac accatttgcac tggcgttgc gggacacatg aagactctgt 138300
ctcaaaaaat aataatTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT 138360
aaagatgtatg ttatTTTTTCA gggatgttttca atttatttttca gacccatcaat aatgttttttgc 138420
aatttgcgttca ttttttttttgc ctgttttttttgc ttgtatgtat aatgttttttgc tatatgtttt 138480
tgttttttttgc agacccggatc tggcgttgc gggccgggg 138540
gtgcaatctc agctctctgc aaccccttgc tcccttttttgc aacccggatc tggcgttgc 138600
cctcccaatgtt agctggggactt acaggccatgtt gggccggggcc cccggcttattt ttttttttttgc 138660
tttagtagata tggggtttca cccgttgc cggatgttttgc ttttttttttgc ttttttttttgc 138720
tccacccggcc tggcgttgc aatgttttgc ttttttttttgc ttttttttttgc gacccatcaat aatgttttttgc 138780
tgtatTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT 138840
gaaagattat cttcaccagg cgcagttggctt cccacttgc ttttttttttgc ttttttttttgc 138900
ctccggcaggc ggctcaacttgc agttcgaaac cccgttgc gacccatcaat aatgttttttgc 138960
ctactaaaaaa taaatTTTTTCA aatgttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 139020
atcttttctga aatttttttttgc aaggcaatgttca ttttttttttgc ttttttttttgc ttttttttttgc 139080

tctatggaaat	aacttatttag	gaagatatact	ctaaaataaag	atcactttgc	ctaaaataaa	139140
ctgatataatt	gatgttcaca	gaatttttct	ttaaccgac	tgataaaatg	cattattctt	139200
gacgtcaagt	gatccacccct	cctcagccctc	ccaaagtct	gggatcacac	acatgaccca	139260
ccgcacctgg	cattattctt	ataaaagggt	aaatattctag	taaagtttta	tgccctttt	139320
gttcatgtac	cattgtttat	tttcttccct	tcctactcac	agtaatcatt	ctttaggtat	139380
gcacttttgt	ttgttattttt	ttatgttaatt	gatattacgc	tccattctgt	acgttgtaact	139440
ttcattcaca	gtgagttttt	gacattccta	tgttcatcta	tacagactta	cttcattttt	139500
actacactgt	agtattccgt	atgttaattt	tactataact	catactgtt	gcagagcatc	139560
tcatagtgta	tgttattactg	ttttgcatt	ttggtatcaa	tgatgttta	agtcatgttgc	139620
agttttttcc	ttttataccc	agtattacag	aggatctt	tttatatgtot	tctttgttacc	139680
aaggaggcaga	ttaaaaaattt	ttttttgaa	aaaattttt	aaaaaaaaatg	aatgaagtc	139740
tcactatgtt	gccccaggctg	gtctcaaact	cctaggctca	agcaatcctt	ccatcttggc	139800
ctcccaaagt	gctgggggta	caggcatgag	ccaccatgcc	ttgcctacat	tttaaatttt	139860
gatagcttt	acaatttact	ttgttaaagta	tctgcatcat	tttatgttct	caccagtctt	139920
taataaagaat	actttcatact	tttggctgga	cacaggggt	cacgcctgtt	atccccagcac	139980
tttggggagc	cgaggcgggc	agatcaagag	atcgagacca	ccctggccaa	tatggtaaaa	140040
ccctgtctct	actaaaaata	caaaaattag	ctggggctgg	ttgcgcaccc	gtagttccag	140100
ctactcgaga	ggctgagaca	ggagaatcac	ttgaacccgg	gaggtggagg	ttcagtgaa	140160
cttagatcac	accactgcac	tccagcctag	caacagatgt	agactctgtc	tcaaaaaaaaa	140220
aaaagaatac	ttcagactta	atttttttt	cagtttaag	tgtttgtta	ttagatttgag	140280
tttcttttgg	tatgtcttt	gatttttctag	tttttttctt	ttatgtattt	actgttcatc	140340
tctttttcac	attattttctg	ttgggttatt	ttattttgtt	tttgcattttt	ttctctcgata	140400
tttttcagca	tgacacttca	ttattcaaaa	aaaaaaaaag	attctctatg	tttctttttt	140460
ctaattcattg	gttggtaata	cctttttttt	aagaccctta	ctgttattttt	tgcttttttt	140520
tttttttttt	tttttttttt	tttggatag	agttttgc	tttttttttt	gtctggagtgc	140580
aatggatgt	tctcggtct	cagctactgt	caactgcac	cttctacccct	ctgttttcaag	140640
caattctctt	gccttagct	cccaagtagc	ttgggattaca	ggcatccacc	accacaccca	140700
gctaattttt	gtattttttg	tagagacagg	gtttcaccat	tttggccagg	ctggctctaa	140760
actactggcc	tcaagtgtatc	cgcctgcctc	ggcatcccaa	agacttggtt	ttacaggcat	140820
gagccacagt	gcctagccac	ttttttgtt	ttaactttgt	ttatagtttt	tatagtttttta	140880
gtataaaacag	atgtatgtat	acacacaact	atggctttat	atatgtttt	agtattttttt	140940
agagcaaggc	tcaccccttt	ggtgcttttt	ttacaaaattt	gtcttggctt	ttcttggcc	141000
tttttttctt	tttggtaattt	tttagattgt	gaatttacctg	ttgactcacc	atgttttgtt	141060
aactgaggat	tttggatggaa	attgcactca	attaaagattt	atcttgcattt	ctgtgcagca	141120
atgtttttat	tcaaaaataatc	cctacttttt	attactttag	atagctataaa	attgtttttt	141180
tggctttctt	gatttagatg	aaacgcgtt	aatttatttt	tttcttctta	atttttttttt	141240
gattgtttaga	agtttaatgtc	ttctgttcat	tcttattttt	gaagatgaca	tttggaaagag	141300
tctagtactt	ggggcaattt	atccgagaat	ctgagcctga	acctgtatgt	aggaatcaa	141360
aaggtttgg	gtgtttttat	acttcataat	aagcctttac	tcacattttt	gattgtactgt	141420
aagtcaaaaga	ccacttaagg	tttttttttt	ttttttttttt	ttatagttttt	ttttttttttt	141480
acccgtgttt	tatagtccaga	aaacgcgtt	tttttttttt	ttttttttttt	ttttttttttt	141540
tctccctctg	atcaaaaatgt	ttctgttcat	tttttttttt	ttttttttttt	ttttttttttt	141600
aaaaaaaggat	gtcttattttc	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	141660
gtttgagacc	gttttatttt	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	141720
catgttcttg	tctgtcaccc	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	141780
tctgtttttt	aggttcaaggc	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	141840
gtgcgcacca	cccacccctgg	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	141900
ttggcttagcc	tgttctctata	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	141960
gtgctggat	tacaagtgtt	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142020
ttccctcaaga	ttttagaaag	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142080
actggatata	gaaaatgtgg	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142140
tgagttctatg	tccttttttg	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142200
gtctcaagga	tagaaaaacca	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142260
gagaacacat	ggacacaggg	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142320
gctgggggag	gaatagcatt	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142380
gcacaccaac	atggtacatg	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142440
ctagaactta	aagtataata	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142500
tcaaaaatca	aattttttat	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142560
aacacaaaat	ttttttgtca	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142620
acacagggaaag	atttcagaaaa	tttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	142680

ttttgggtga attgacagta taaaagatac taatttttgt aagaagaaaa 142740
attcaatttt gataagtatg ttaagatta agacttattg gccaggcgct gtggcata 142800
cctgtatcc tagactttg gaaagctgga gcagggtggc caccaggctc agagatttag 142860
accatccctgg ccaacatggt gaaaccctgt ctctactaaa tttagccagc gtggggcac 142920
atgecctgtc accccccttc gggtttaagc gatctactg ctcaggcctc tgagtagct 142980
gggattacag ggcgcattgc tggccaggtt ggtctgttct caaactcttg acctcagggt atctgcccgc 143040
tggccaggtt ggtctgttct caaactcttg acctcagggt atctgcccgc 143100
caaagtgtcg ggattacagg tttttttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143160
gtgcgtatcc agctcaetgc ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143220
tcttccaagt agctgggatt tagtagagac agggtttcac ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143280
tctgccccacc tcggcctccc atcttatttt ctttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143340
caggcttgc tcactgcaac ccaagttagc tgggattata agtagagatg ggggtttctt ctttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143400
cccaagttagc tgggattata agtagagatg ggggtttctt ctttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143460
cctcccaaaag tactggctt ggctgcattgg ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143520
ggctgcattgg tggttcatac gtttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143580
gttggagggcc aggagggtga ccaccttagt aatgggacaa atactggaa gaggtcagtgc ggcattggcc aggaagaact 143640
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143700
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143760
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143820
atactggaa gaggtcagtgc ggcattggcc aggaagaact 143880
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 143940
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144000
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144060
atactggaa gaggtcagtgc ggcattggcc aggaagaact 144120
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144180
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144240
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144300
atactggaa gaggtcagtgc ggcattggcc aggaagaact 144360
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144420
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144480
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144540
atactggaa gaggtcagtgc ggcattggcc aggaagaact 144600
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144660
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144720
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144780
atactggaa gaggtcagtgc ggcattggcc aggaagaact 144840
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144900
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 144960
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145020
atactggaa gaggtcagtgc ggcattggcc aggaagaact 145080
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145140
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145200
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145260
atactggaa gaggtcagtgc ggcattggcc aggaagaact 145320
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145380
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145440
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145500
atactggaa gaggtcagtgc ggcattggcc aggaagaact 145560
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145620
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145680
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145740
atactggaa gaggtcagtgc ggcattggcc aggaagaact 145800
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145860
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145920
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 145980
atactggaa gaggtcagtgc ggcattggcc aggaagaact 146040
ggcattggcc aggaagaact ggggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 146100
gggtcttc cccagtattttt ttttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 146160
tggttttttt tcttgagatg atcttgcctc actgcaactg caaagttagc gggatttttttta gacggaggt tgcgtgtcg cccagagctg gagtgcata 146220
atactggaa gaggtcagtgc ggcattggcc aggaagaact 146280

57/122

agtattaca	ttagagatga	gaagaaaagag	gtgccttta	cttttcaata	taccccttcc	146340
tctgttttt	gaactttctt	gccctatgca	tacgttattg	cttaatcatc	caccctcatct	146400
cttccccgt	ggcttctgt	tgcatttga	atgaaatcta	gccttcttgc	tgttacctgt	146460
ggatgtccct	tgtggccctc	tatcaccca	cttggaaacca	ctecccatt	ggactgagct	146520
cttcattgtac	tatctttttt	tctttgtct	aaggttctt	actttgaatg	cctctgagct	146580
tgttatttca	tggctgtggc	aaggccctgc	atggcttca	tgcaggatg	gttccctcctt	146640
ctcatctcaa	tattatctct	tcagagaggg	accttccaa	ctcccgatgat	ctaaaatcct	146700
ttgttatatac	cactcaactac	caacttcttc	tttttttttc	cttttatctt	tttttttttt	146760
ttttttttt	gagatagggt	cttgctctgt	tgcccgagct	ggaatcaacga	ctcaactcgac	146820
cctcatcttc	ttggctctcaa	atgatccctt	cacctcagcc	tctcgagatg	ctggaaactgc	146880
aggcacacac	caccatactt	ggcttattat	tttacttttt	gttagagacag	ggtttaccca	146940
aggctggct	caagctccctg	ccgcaagcaa	tccacatctc	tcagccccc	aaagtattgg	147000
gattatagga	gtgagccact	actcctggcc	tattttctta	ttcaactgtct	aaaattatct	147060
tgttcatta	tttacataact	tgtttagatc	ttttttctca	gctggacatg	gtgcccacca	147120
cctgttaatct	caatactttg	ggagggctgg	ttggagaat	gtttagagccc	aggacttcaa	147180
gaccagctg	ggcaacaaag	ttagacccctg	tttataaaaa	atttttttaa	aattagctgg	147240
gcatggggc	acatgcctgt	ggtcccagct	actttggagg	cagaggtggg	agaatcgctt	147300
ggcccccagga	ggttgaggcg	acggtgagcc	atgattgtgc	cactgcaactc	tagcttagtg	147360
acagagtgg	accatgtgtc	aaaaaagtaa	ataaaaatag	tttctcttcc	atgactagaa	147420
tattacctt	atgtgggcag	ggaggttgc	tatactatt	ggcactatat	ttctgtatc	147480
tgaaattatg	ccttagacat	ggtaagtaatc	ccttaaatat	ttattgactg	aattttttaa	147540
tacttaagaa	tttcattttgg	gattatctga	gtggtaagat	tacggattat	attttatgtaa	147600
aaaaaaatca	ttttttaaac	ttggttgccc	tttgccacac	tgacatagac	actaagtttt	147660
cttagccaga	ttacttccga	ggataactcac	agaggccatt	ctcttctcaa	tccccaataa	147720
attgtatatt	tttagactt	tcaagtcataat	gcaacttta	gatgtatgtat	ctgtgtat	147780
catatccca	tttcataaaat	gttagaaattt	aagtctgggc	acatgggtc	tcacccatgtaa	147840
tctcagcagt	ttggggaggcc	aaggcgagcg	gatcaactgag	gacaagagat	aaaggacagcc	147900
tggcaacat	ggtaaaggct	tgcctctatt	aaaaatacaa	caatttagggc	cgggcgtgtt	147960
ggctcacgcc	tataatccca	gcacgttggg	aggccaaggc	aggcagatca	cgaggtcagg	148020
agtcgagac	cattctggct	aacacagtga	aaccccatct	ctactaaaaa	tacaaaaaaat	148080
tagccagca	tggggcagc	cgcttgtatg	cccagctatc	ggggaggctga	ggcagggtaa	148140
tcccttggaa	ccggggaggcg	gagggttca	tgagctgaga	ttgcaccgct	gaactccacg	148200
ctggtcaaca	gagggagact	ctgtctcaaa	aaaaaaaaaa	aaaaacaattt	agccaggcgt	148260
ggtggcgggt	acgatgtact	gtaatccca	ctactaggga	ggctgagggga	ggagaatcac	148320
ttaaaccagg	gagggtggagt	ttgcagcggg	ctgataatgc	accactacat	tccagccctgg	148380
gcaacagat	gagactctgt	ctttaaaaaaa	aaaaaaaagaa	agaaaaagaaat	tgaggaatgt	148440
ggagattgt	gtctgtatt	tggttaggaat	cacacagcag	tttagtagca	actacagggc	148500
tttggttcag	aataccacat	tgacaatgtt	ttgtttacag	ttcggtctcc	cttccctctgc	148560
ctttctctcc	ttcatttattt	agggcagctg	gaaagaattt	tcatcatatta	ctagcctata	148620
gctttaattt	gagttttgaa	accttgataa	tagacacag	agaaaaaaagac	tgagttttct	148680
ttttttgaga	cagtcttgc	ctatggccca	ggctggagt	cagtgcacacc	atctcagctg	148740
gttgcacact	ctgccttcca	ggttcaagca	atttcgcctc	agcctctcga	tgagctgaga	148800
ttacaggcac	gtgtcaccac	gcccagctaa	tttttgcattt	ttgtttcggt	ttgttttttt	148860
ctgagatgga	gtcttgcct	gtcaccctagg	ctggagtgc	gtgttgcgtat	gttggctcac	148920
tcaaaccctt	gtctcctggg	ttcaagcaat	tcttctgcct	cagcctcccc	agtagctggg	148980
actacaggtt	cgtgccccca	tcccttagttc	atttttgtat	tttagtagta	gatgggggtt	149040
caactatgtt	accaggctgg	tctcgaatctc	ctgatctcag	gtgatctact	cgtctcagtt	149100
tcccaaaatg	ctgggattat	ttggcacacgc	ctatttttgt	attttttagta	gagacgggggt	149160
ttcaccatgt	tggtagact	ggtctcaaac	ttctgacccct	aagtgatttg	cccgccccag	149220
cctcccaaaag	tgctgggatt	acaggcgtga	gccaccgtgc	ccagccaaga	tttagttttg	149280
aaaagaggct	tctgagatta	tgagaaggcc	aagcaagata	acttaagaag	ttacattaaaa	149340
atcatcttaag	agacagtgtt	acaagaaggaa	attgtttaaaat	gatgttatga	gcacgtcccc	149400
aatgtatgtt	caatccccctg	tgcttcgata	cattgggtgg	agacaaaact	gtacttaaat	149460
tgataaaatcc	tttacatgtc	attttaagga	gcttagactgt	actcccatca	tgtagacatc	149520
agagattttt	tttttttttt	tttttttttt	tttttttttt	tttggacacag	agtttgctc	149580
ttgttgcgca	ggctggagtg	caatggcggt	atctcggtc	accacaaccc	ccacccctcca	149640
ggttaaca	atttcctgc	ctcagccccc	cgagtagctg	ggatccacgc	catgcaccac	149700
cacgcctggc	taattttgtt	tttttttagtag	agacgggggtt	tctccatgtt	gtggctggc	149760
tcgaacttct	gacccatgtt	gatccctcccg	ccctcgaccac	ccaaagtctt	gaaaatcag	149820
gcgtgagcca	ccggccccag	cccagagatt	tctaaaacaga	gttctaacc	gatgttttc	149880

cctgtcagta	gaatgagaat	gaattggagg	tggagagac	tggcatgagg	gacaccagtc	149940
agccagtgga	attagctgg	aatgttgata	ggagaagaaa	aagattcaaa	gttaggtagt	150000
ggtagcaaga	attagaggg	aggtcgatt	tatgatatgt	ccaaggttga	attctaaggt	150060
gaaatttttgt	ggcagattt	atgtgtaaat	tggaaaggt	gattgagttt	tttaacatg	150120
ggttttctaa	catgtcaata	gagtgactct	gcaggggggc	ctgacgagag	aacagtgc	150180
gggggttattc	aacagccagt	tgagcattca	tgcagagcat	ttaacactgt	gactctgtag	150240
actctggttt	gcgataaaat	ttccatcaa	caatattaa	acccttaggt	aataataaaa	150300
attgaggaa	aaggatccag	gttttgtatt	tttatgaat	tcagttatgt	aattaaacag	150360
gaccttgcct	caagaataaa	tctaccaaca	attaacttgc	tttaaagcaa	agtaggaag	150420
tgagcatgtt	caaattatta	aataaaaaag	taagctgtgt	atttcattca	tagaaataga	150480
ggctggccta	cttcggatga	ttctcagcat	gtgattacag	atgtggctt	atacatccct	150540
gggagtttaag	gcgtactctg	gcttggatag	agtagagctc	tttgaactc	ttctctcacc	150600
cagctatgtt	atataactgt	gagaactaga	atagtagcgc	atactctgc	tttagaagccc	150660
ttttatata	gagctggct	ggaaggttt	aaaacataac	aatatgttt	gtgtttccca	150720
atgttattgt	agattcttac	ccaagagcat	tatcctgggt	agggttgg	ttggttttgt	150780
tttggttttt	aatgtttgcc	acaaactaac	actagatgtt	agttcttca	tcaagtgg	150840
agagtagaa	aaaagtccag	aactctgaaa	cacccatttca	aaagttttt	aagccatgtat	150900
gtttcaagt	ttttagtct	gttatgaa	caatataatc	agtttttatt	aatgttaacat	150960
tccttagtgt	tttgggtat	cacacaaaaa	agaatattca	tatctgaaag	caacagctt	151020
taaataagag	cattgtgggt	gtgggtgt	tagtggttt	ttttttttt	tttgagttgg	151080
agtctcgctc	tgttgcgg	gttggagtgc	agtggcacga	tctcagctcg	cttcaacctc	151140
tgctccagg	ttcaagcaat	tcttcgct	cagcctctg	agtagctgg	attataggca	151200
cctgtacca	tgcctggct	attttttata	ttttagtga	gacagggttc	accatgttgg	151260
ccaggctgg	tttgaactct	taacccagg	tgaatcaccc	acccctggct	cccaaaagtgc	151320
tggaaattaca	ggcatgaacc	accatggcca	gccaataaaag	agcattttt	atgtaaaatt	151380
atgcataaaa	tgtacattca	attttgtt	tgttactag	gatccatgtt	ctcacaagct	151440
atgaagaaat	gggtgcaagg	aaatactgt	gaggtaataatc	ctacccattt	gataaaaaaaga	151500
tttctgttta	taatgtccac	cctcatgtaa	tgtaggttta	aaattttctt	tttcttttagg	151560
tcccatgttt	aagcagcat	gcacatttat	gttcttttac	ccagaatgt	ccaaagaaagg	151620
gtggccctt	cttaacatct	aacaatttgc	tggtagtagc	agtgaaggt	tcttcagtc	151680
gaggcttagg	ccactgaagg	atatacatgc	attcaagttt	ccatcagcca	gcaggcatca	151740
gtaatcagt	tgtatgttca	aagcttcaat	gttcccttcc	ccacttgcag	tttacttca	151800
agtagtggag	gtttgttctt	ttaatagtta	attaagtaca	ttgagatgt	ggaggtgaaa	151860
aaaggaaaat	ttttttttttt	gaccatctaa	tgatggaaatgt	gttccgtt	aggtatccag	151920
tagttgacac	ttggaaagacag	ggaatgacat	gttataattt	atagccagag	gttggcccg	151980
gttttttgt	acatggaaat	gaaatttta	tccaaataag	tagaaattat	gtgcgtaa	152040
cattttttaa	gagcaactgag	tatgtgcattc	togatccatc	taatgataaa	ccatttatcac	152100
cagttaaat	tattttttttt	aggccaggaa	agagcttagt	tggatgtt	ctaaaatgtat	152160
agtcgtgtac	attatgcagc	aggctcgtat	tgtcaacc	ttagagaaat	ctacaaaggt	152220
aaggatgact	tgcgttttgt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	152280
agaacataaa	gggtttctt	gcctttaag	gattaactgc	tgtggggatt	acccctttat	152340
cataagcaac	tagaaaattt	acaaactaaa	tggaaacaat	gtttgcata	attggacaat	152400
gggcaataca	ggggaaaccat	ggggaaaccat	cgagccccag	tagtctgt	gaacgaaaga	152460
gttaaatata	aaagttcagg	ccagggtcag	ttgtctcacgc	ctgtatccc	agcactttgg	152520
gaggccaa	cggggtgaatc	acttggatgc	aggagttca	gaccaggctg	gccaacatcg	152580
tgaaaaccct	tcttagccgg	gtgtgggt	aggcacctgt	aatcccaact	atttgggagg	152640
ctgaggcagg	agaatcggt	gaaccaggga	ggcggagggt	gcagtgagcc	gagatcacac	152700
cactgcact	cggctgggg	gacgagcgc	accccatttc	aaaaaaaaaa	tcaatgttca	152760
gagagctaa	ttttagttaga	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	152820
aaagaaaagcc	gttagagat	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	152880
gtatatgtt	gggggtaaaa	cgcatgtgc	caggtagaga	accccccaga	aatttagtgg	152940
ctgaatgatt	gttggaaat	agggctttaa	aaagttcat	gccagaagga	tctggccaga	153000
gttagagagac	tttagtaata	acaaggctt	gggtgtgt	tttcacagg	ttatgtcctt	153060
ctactgaata	taaatttagt	ctagatgt	acaacctgt	ccaagttca	aagcaaattt	153120
ttaaagggtt	aaatttactt	acaactgtat	gccaatcaaa	aggccttacc	ctctttacag	153180
taacacaa	aaatttcagca	cttcacagt	taaagtttta	atgtctgac	tccaggctgg	153240
gcgcagtggc	tcatgcctgt	aatcccagca	cttggggagg	ccgaggccagg	tagatgac	153300
gaggctgg	gttcaagacc	agccctggct	acatggtca	accccgcttc	tataaaaat	153360
acaaaaactt	agccaggcat	ggtggccggc	actgtgtatc	ccggctactt	gggaggctga	153420
ggcaggagaa	ttgcctgaac	ccaggagg	aagggttgcag	tgagccgaga	tcgcaccact	153480

59/122

gcactctggc	ctggggcaaaa	agagcaaaac	tcaggctcaa	aaaaaaaaaa	gaatgtctga	153540
cgtcaatcac	aaatttacca	gcatgacatg	aagtgtac	ataaccagga	gaaaactcaa	153600
tctatagaaa	cagaccggaa	tgtgagaaaag	atgtatgaa	tagcagacaa	agaccatcaa	153660
gtggctattt	taaatattaa	aaatatgtt	aagtggccag	gtgcagtggc	tcatgcctgt	153720
aatcccggca	cttggggagg	ccaagggtgg	taggagttca	agaccagctt	ggccaatatg	153780
gtgaaacccc	ttctctacta	aaaatacaca	aaaatttagct	gggcattgg	gcagggtgcct	153840
atagccccag	ctatatggg	ggctgaggca	caagaatcac	ttgaacccgg	gagggtggagg	153900
ttgaggttgc	agtaagccg	gattgtgcc	cttgtactcc	agcctggaca	acagagttag	153960
actctgtctc	aaaaaaaaaa	aaaaaaatgt	taaagaaaac	aagagtataa	tgagaaaaat	154020
gcaaaaatagt	tttaaaagaa	ccaaatggaa	tttctttaaa	aaaaaaaaat	cagaatggg	154080
ggccggcg	ggtagctcac	gtctataatc	ccagca	gtgggggtcg	aggcaggcag	154140
atcacctgag	atccgttagt	caaggccage	ctgaccaaca	tggagaaaacc	tcatctctac	154200
taaaaataca	aaatttagctg	ggcgtgg	cgcatggct	gtaatccca	ctacttggga	154260
ggctgaggca	ggagaatgtc	ttgaacccgg	gaggcagagg	ttgcgg	ctgagattgc	154320
accatgcac	tccagcttgg	gccaca	tgaaactcc	tctc	aaaaacaaaaaa	154380
aaaacagtag	actcgaagaa	ctagctgag	tttctttac	tttaggc	aagtgtgacc	154440
tttgcaggt	gactacttta	gttccctatg	tcctcattag	tagatcag	aaattcgaca	154500
ccaaaacccc	aaaagaaaaaa	ccccttctaa	tcctcattcc	atgattttat	gaatgcata	154560
agtccctaggc	ctcgcaagga	atacttcat	tcttatcc	gtgttgcata	ctctctgtct	154620
caacccctaa	ctcgacattt	gcctatagga	tgtacttgg	cattcag	aaactaccc	154680
acaccattac	tgaatttgc	catgtgcaca	tgtccatgc	cacaatacc	gggaccc	154740
cttccgtat	atttgcgc	agtgcgtg	ctacaggagg	gagtcgt	atgtctgc	154800
gtgtgtctt	accatccctc	ttgaatatgc	tctagggtt	attcctagaa	gtagaattac	154860
tctattgaaa	attggcaata	ttttcattc	taatatctat	tgccaa	catggaaagcaag	154920
tctggatg	actccttgc	atatgc	tggtaagtt	acgtaa	acttcc	154980
tgttca	tat	ttttaaca	aggaaaaat	tttgc	tttagtgc	155040
cttctggctg	tcttaaactc	tggtatata	taaacactaa	gtgttgc	ccatcc	155100
tttgcata	taa	ttgtcactt	tttagagaaa	tttgc	ttttctt	155160
tttcagttat	gactcaaaac	ttgagataaa	ggaaatctgc	tttgaa	aaaa taagagaact	155220
ttttccctt	gggtggat	ttcaacacag	ccatgaaaaa	cagcacta	tttgc	155280
gtcactgtt	tttccaggag	agaatgggag	acaatcc	acttcc	cc	155340
acctgttaggc	ataatttgc	cacatgtat	tcacac	agagtctt	aa	155400
tggatgtt	tacattacta	gaaaattatt	agtttccaa	tggcaataa	ccatttat	155460
gagtgttta	gcctactgg	atagacagg	accacatcc	ctggga	agca gata	155520
gaactgatac	ttgatgcaca	ctcgtagtgg	taactcatcc	ctaata	cgca ttgaa	155580
ggtgcac	ggtgttgc	ttgtcctcc	aaagcagg	atgc	ccccc accgag	155640
aggcagctt	gagtggc	gtgggtct	cagttc	ggaa	acgggt gagagtt	155700
tatgcactt	tcttgcac	ggcattaatt	ttgaaaggaa	ctgaca	atgggtca	155760
tttcagtgac	tttcccttcc	ctctgtatgc	agtatata	tttgcattt	taattcctcc	155820
tcctgagat	cacttactt	aaaaccattt	tctccctgc	taacaga	agg gtgt	155880
ggtttactt	gagcattagg	atttgc	ttgaaatttct	gcactcc	gtacttaactt	155940
tcccttcaga	atacatgtt	aaagaaaaaa	agaaatagcg	atgactcc	ttttgc	156000
gtggcac	taa	gaaacaagca	gttcttcc	ttttttttt	taataagg	156060
gaggcaggat	actggggaga	gagaaacatt	tgacttgc	ttgc	cccc atttttgc	156120
gtgagctgg	aaatgtgc	ttggcgt	tttttcttct	tttttttgc	tttttttgc	156180
ctcacttact	gcacttctgc	tcagttgc	tctgc	tttttttgc	tttttttgc	156240
gactgcccag	ttggtgagca	cactccattt	accacgtgg	tttttttgc	tttttttgc	156300
catgatttgc	aggaaagaaa	gttctttag	atgttactgc	tttttttgc	tttttttgc	156360
aaaaaaaata	tatatatata	tgtataaata	tataattt	aatcactt	tttgc	156420
aagtcttgc	tgaaacagaga	atttatttca	ttgcaatatt	tttttttgc	tttttttgc	156480
gtttcatcg	cagaagaagc	aaaaagg	tgtgtaa	tttttttgc	tttttttgc	156540
ctgttatttct	tttaaagctg	aagtatttcat	gtactttaac	tttttttgc	tttttttgc	156600
ttgattttaa	aatatata	tatgaatttct	atttaaaatt	tttttttgc	tttttttgc	156660
gggcattat	ggcttcttgc	ttgaaatata	ttgatcttc	tttttttgc	tttttttgc	156720
ctaaaaacc	agaacatgc	ccactactgg	actttgc	tttttttgc	tttttttgc	156780
aaacccaa	tttttatttgc	tcatctatgc	tgtgat	tttttttgc	tttttttgc	156840
aatatttcc	tccacttcc	attgttccaa	tcttttacag	tttttttgc	tttttttgc	156900
agaataaaat	gtgggtttgg	gcaacttgc	ctgcctctgc	tttttttgc	tttttttgc	156960
ctacaagact	ctaggctgtt	taaacttagtgc	tttgc	tttttttgc	tttttttgc	157020
tttgatata	cattttgtgc	ttctgagctt	gagatgc	tttttttgc	tttttttgc	157080

aagagaatag	cagcaaattt	gagactcggc	tactttttc	tgccccacct	gctttgagac	157140
acagaagcgg	agtgtggcc	gaaattatta	gccagagttt	atatttgatc	taaagttaggt	157200
ccttgactc	attttaaagt	tggaaatttga	ttcctccaac	attgagcacc	caccatgttc	157260
caggctctgt	gcattgtgcc	cacaaaataa	gattccctgg	tggagtttt	atgggttcaa	157320
ataatcgtt	gaacacccctt	catcttatac	atgttgttga	cattgacaca	aattgtttaa	157380
aaagaaaaaga	tatttagagag	aaagtggta	cttggtaact	tgatgtgtct	tcatcattcg	157440
gtaagattt	atgaaagtaa	aaagcaaatg	tcagccaaat	ccagtgaaca	gcaataaaac	157500
agggagtaac	tttttataac	tttttctact	tggatttca	cattcagtag	agcttttgcg	157560
aatagaatg	gttacatgt	ctggaggtt	gactagtta	gttaggattt	ggaggggaag	157620
gtcattctga	attgtaaaca	agtacaaact	tctttgtgt	tttatttaag	tactgagagc	157680
taagcacctg	atgaagtgac	tgacctctct	ccagtgacag	tgtttggta	cctgcctgac	157740
ttcaggagtg	gggtttatgt	ttctacacag	tgaccttttc	tctcgccctc	tctccctct	157800
tgcccacaca	ccagttgtt	ggacccctgg	tgaactcttc	atccagacag	gcccaagaca	157860
gttcttaatg	ttaagaat	tggggccgg	cacgggtct	catgcctgt	attgcaacac	157920
tttgggaggc	cgagacaggc	ggatcacttg	aggtcagggg	ttcgaggcga	gcctggccaa	157980
catggtaaa	ccctgtcttt	actaaaaata	caaaaattag	ctgggcatgg	tggcgacacg	158040
ctgtaatccc	agctacgtgg	gtggctgaga	caggggaatc	gcttgaacct	ggaggcggag	158100
gttgcata	gagccgagac	cgtgtcactg	cattccagcc	tgggtgacag	aggagactc	158160
tgtctccaaa	aataaaaaata	agaaaaaaaga	ttttggctta	gttgcagtgg	ctcacgcctg	158220
taattacagc	attttggaa	gccccaaatg	ggcagatcac	ttgaggacag	gagttcgaga	158280
ccagccttga	caacatggt	aaactccatc	tctactaaaa	agacaaaatg	tagccagatg	158340
tggtgatggg	cacctataat	cctagctct	cgggaggctg	gggcaggaga	atcacttgaa	158400
cccaggaagc	agagattgca	gtgaggccaag	atcacatctc	tgcactccag	cctggcaaac	158460
agagcaagac	tctgtctcaa	aaaaaaaaaa	attttggccag	gcccgtt	tcacgcctg	158520
aatccagac	tttggggagg	ccaaggcagg	cagatcacga	gttcaggaga	tcgagattgt	158580
cctggctaac	atggtgaaac	cctgtctcta	ctaaaaatac	aaaacattag	ccgggtgtgg	158640
tgggggcac	ctgttagtccc	agctactagg	gaggctgagg	cagaggaagg	atgtgaaccc	158700
aggaggcgg	gtttgcaga	agccaaatgc	gtccactgc	actacatct	gggcgacaga	158760
gtgagactcc	gtctcaaaaa	aaaaaaagaat	tttggccgg	tgcggtggca	catgcctgt	158820
gtcccgac	tttgggagac	caaagtgggc	ggattacctg	aggtcaggag	ttcaagacca	158880
gtccggccaa	tatggcggaa	ccctgtctct	tactaaaaaa	aataaaaaaa	ttagccaggt	158940
gtggggcgg	gcacctgggg	aggctgaggc	agggagaaat	gcttgaaccc	gggaggcaga	159000
gtttgcagta	agccaaatgc	gtgcaactgc	actccagac	aagactctt	ctcaaaaaaaaa	159060
aaaaaaaaaaag	aattttgcat	ggggaaaggag	agatactgtt	caccatctgg	aatggtgctt	159120
ggatgtggc	tttacaaaaat	caggagccag	cactcatgg	aaaaacacaa	gcatgtggc	159180
ctgagatgc	aggtaatctg	ataaccctga	agacatcctt	ggtttctgca	tctatttcctg	159240
catcctgca	ttggactaca	ttaatctgtc	agtttatectt	ataatgattt	tttatttttt	159300
tttttgaga	tggagtttgc	ctcttgcgtc	ccaggctgga	gtgcaatggc	acgatctcg	159360
ctcaccacaa	cctccacccct	ccaggttcaa	gttattctgc	tgcctcagcc	tcctgagtaa	159420
ctgggattac	aggcatgcgc	caccacact	ggtcaatttt	gttatttttg	tagagacggg	159480
gttttccat	gttggtcagg	ctggctctga	actcccaacc	tcaggtgatc	accctgtctc	159540
ggcctcccaa	agtgtggg	ttacaggcgt	aagccatgg	accggcttg	ttttttgatt	159600
ttttgaaacc	agtctgaagt	gagtttttt	aattacgtga	aaggagttt	gtctaaaatac	159660
tgccatactg	ccctatgtcc	taatgattt	gtattctca	catgtctgc	aatgtactgt	159720
gattttctgg	gaataatttt	tcttttagta	attttactca	agtcgtcatg	tttattctct	159780
aaaaatggta	tcctaaccct	atggagctaa	aagacacccc	ttgtttttat	aacaaggcgt	159840
tactgaggcc	caggaagggg	agaagtccct	ggcttgcgg	atgatcacca	ttagaactca	159900
ggcctggggcc	agtgccttt	catgcctctc	agatccttc	aaagaataat	gaagattata	159960
accgcctttt	gcaattgtaa	taaaccagg	aatagaaagg	ttttttgtt	gatgtactgg	160020
agaagtttgg	tttttttatac	aaattttgtaa	atacctgcctt	tttttttttt	tttttttttt	160080
ctaggcaagg	tctctggcct	tgtttttttt	ctcaagggtt	caactttgtt	ggccccacact	160140
aatagttacc	cactgaggcc	ctctccgggt	gaacatttgg	cacttagagga	agccccctctg	160200
cttggggcagg	actgggcgt	gtgcagatgt	ggagcggtga	tactgtggat	tctggggcagg	160260
tggagatggc	cagtgtatgc	caataaaagg	cactgggggg	agcagttgt	gttttttttt	160320
tgaggggcatt	catgttcagg	gggggttgc	ggccactggc	ttgttttttt	cacaggagag	160380
tgggtattcc	tgcccttagta	actttatgtt	aacaaggattt	tcctcgttct	gttccctctca	160440
aactgcctgc	tctggcata	tcagaatgtc	acagaactca	cctggatgtca	ttcagccccct	160500
tgcctaaagg	tgacagtgtc	tctccctccc	caccccaccc	ctcataccac	tgaagcaccct	160560
gtcaagactgg	cccagtctgt	gggcaaggag	cctagagagg	gctttagttt	agcttggaaag	160620
gagctgggat	ttaccaagaa	gcaaatgaga	gacgaggatt	gcaacaactg	tgccttcc	160680

61/122

ccagcttcag	ctgactcctg	tatattgact	gtgccttcag	actcatccgt	aagtgacccc	160740
aggctggcct	ctccccacatc	acagtaagaa	ttccacacac	cataacaat	ggaagaggc	160800
tccagctgaa	ggaagcccca	cacttcttc	aagttttct	tagtcttctc	ttcttggcaa	160860
agagtacctt	ttgtttcttc	taattatgt	actattggtt	tagtaaatat	tcacccattc	160920
agtcaaccctg	taagtggcag	gcactgtta	cagggacaca	ggaaggaata	aaaacttgca	160980
ggcaccttgg	agcttgcat	ctattgaaga	ggttaatggaa	gttgggatag	cagctaaact	161040
atgctggtat	tggccaggcg	cagtggctca	cacctgtat	cccagcaact	tggaggccaa	161100
ggtggcaga	tcatgaagtc	aggagatcg	gaccatctg	gctaacatgg	tgaaaccccg	161160
tctctactaa	agtaaaaaaa	aaaaattagc	caggtgttgt	ggcgccgc	tgttagtccca	161220
gtctactgg	aggctgaggc	aggagaatgg	tgtgaaccca	ggagggcgaag	attgcagtga	161280
gccgagatgg	caccactgca	ctccagcctg	ggtgacagag	cgagactctg	tctcagaaaa	161340
aaaaaaatatg	ctggtagttt	tgattcaaga	tggccttgg	agcccatgtat	ttaggtctcg	161400
taccaccaa	ggtctactgg	aaaacatcg	gctctctgc	tatagaccca	tagggagac	161460
tgagccgag	agggggagct	gaagagaatg	gccccttctg	tgtcctgtca	gcctcatct	161520
tccgcaaggg	ccaggctgtc	tgccacttca	ttcacttgct	gcaagactgg	aggttttcc	161580
tcaggtttg	agcacctgtt	ttacaagatg	tcagcatctt	gatgcctgag	accatcaagg	161640
caagtctctg	aacagggctt	accttagatg	aaggcttaga	agaggccgt	aagtcaagtct	161700
cagctccgtg	gctctgcaga	gctttggac	atgtgaattc	ttaaaaacaa	gactattgt	161760
cagtactat	atgcatgcag	tataaaattt	taaccttgg	aaatcttagc	tagtgttga	161820
gctaatttca	taaaggaaatc	agctcttgag	ttctgcagtg	gtaataataa	tcagcataat	161880
gatgaaacac	tgtgtgtgcc	aggcagcgtc	tcatttgatc	cttgcataaa	tctgttaagt	161940
actgattttc	tcccttctt	aaacaaagtt	ttttttttt	tttttagagag	ggtctcaacta	162000
tggtcccaag	gctagtcttg	aattc				162025

<210> 36
<211> 162025
<212> DNA
<213> Homo Sapien

<220>
<221> mutation
<222> 156,277
<223> Nucleotide Base Change: T to C

<400> 36						
gaattcctat	ttcaaaaagaa	acaaatgggc	caagtatgg	ggctcatacc	tgtatccca	60
gcactttggg	aggccgaggt	gagtgggtca	tgcgggtca	ggagttccag	gcaagtctgg	120
ccaaacatgtt	gaaacactgt	ctctactaaa	aataaaaaaa	ttagccggc	gtggtggcgg	180
gcacctgtaa	tcccatgtac	tcaggaggct	gaggcaggag	aattgtctga	acctgggaga	240
tggagggttgc	agtgagccca	gatcgccca	ctgctctcca	gcctgggtgg	cagagtgaga	300
ctctgtctca	aaaagaaaaca	aagaaataaa	tgaaaacaatt	ttgttccat	atatttcaca	360
aatttggaaat	gtttaaggtt	ttatggtcac	tgatatctcg	tttcatctt	tatataatca	420
ttaagtttga	aatgtataact	tgctactacta	acacagttagt	taatcttagt	cctacaagtt	480
actgcttttta	cacaatataat	tttgcataa	tgtatgcact	ggtgtttatg	tacgtgttta	540
tgtttatatac	tgtttaaaaatt	agcagtttcc	atctttttct	attttgcacc	atcacatcg	600
ttcagaaggg	ttgacagagc	aaaatgattt	gatgaagtat	aaaagtcaca	tggtgagtgg	660
cataaataca	actctgaaca	attaggaggc	tcactattga	ctggaactaa	actgcaagcc	720
agaaagacac	atatacttata	tgtcaagaga	tgtaccaccc	aggcagttaa	agaagggaaag	780
tacacataga	aagcacaatg	gtgaattt	aaaaaattgg	aatttatcg	acactggatt	840
catttgcctc	taaaatcgaa	gtcctctatt	gtttttttgt	ttttgtgggt	ttctttttaa	900
attttttat	tttttgcata	gtcgaggatct	cactgtgtt	cccggtctgg	tctagaactc	960
ctggcctcaa	acaaaacctcc	tgcctcagct	tcccaaagca	ttgggattac	agacatgagc	1020
cactgagccc	agcccagacg	cttttagatt	tatgaaggt	ctgaaaatgt	tgtagaaacc	1080
gcataagctt	tccatgtc	tttcaaaatgt	tgtatgtctc	tttagttaaac	caaccaagtt	1140
atccctcaag	ggcaaaaataa	catttctcg	tgcaaaactg	atgcacttca	ttaccaaaaag	1200
gaaaagacca	caactataga	ggcgtcattt	aaagctgcac	tcttcagagg	ccaaaaaaaaaa	1260
aggtacaaac	acataactaat	ggaacattct	ttagaaggtc	ccccaaatgtt	atgataaaaca	1320
ttttcatcaa	agagaaaaaa	gaacaagggt	ttagcaaaatt	cctctatcaa	ataacactaa	1380
acatcaagga	acatcaatgg	catggcatgt	ggaagggaaa	gtgtctgtc	atgtacaaac	1440
cagtagataa	tttcaactt	ctgcccata	aaaccttctt	gcaaggatgt	aatcagcact	1500

tctcatgttt	gttttgcctt	gttttgcctt	gttttgcctt	gttttgcctt	gttttgcctt	acagggcctt	gctctgtcac	1560
acaggctgga	gtgcagtggc	acgatcagag	ctcaactgcaa	cctgaaactc	ctgggctcaa	1620		
gggatccctcc	tgcccttagcc	tcccagaatg	ctgggactac	aggcccacca	tgcccagcta	1680		
atttttaaa	tttctatag	agatgggatc	tcactagcac	ctttcatgtt	tgatgttcat	1740		
ataacaacgac	caaggtacaa	tgtggaaaag	ggtctcaggg	atctaaatgt	aaggaggacc	1800		
agaaaagaaaa	gggggttgcta	catagagtag	aagaatgtgc	acttcatgcc	agtctacaac	1860		
actgcgtttt	tcctcgagac	agagttgtat	atctaaatgc	ggggtccccca	accccccagtt	1920		
catagcctgt	taggaaccgg	gccacacagc	aggagggtgag	caataggcaa	gcgagcatta	1980		
ccacctgggc	ttcacctccc	gtcagatcg	tgatgtcatt	agattctcat	aggaccatga	2040		
accctattgt	gaactgagca	tgcaagggat	gttaggtttc	cgctctttat	gagactctaa	2100		
tgccggaaga	tctgtcactg	tcttccatca	ccctgagatg	ggaacatcta	gttgcaggaa	2160		
aacaaccta	gggctcccat	tgatgttata	ttacatgttag	ttgtatcatt	atttcattct	2220		
atattacaat	gtataataaa	tagaaaataaa	ggcacaatag	gccaggcgtg	gtggctcaca	2280		
cctgtatcc	cagcacttcg	ggaggccaag	gcaggcggat	cacgaggctc	ggagatcgag	2340		
accatccctgg	ctaaaacggt	gaaaccccg	ctactaaaaaa	ttcaaaaaaa	aattagccgg	2400		
gtgtgggtgt	gggcacctgt	agtcccagct	actcgagagg	ctgaggcagg	agaatgggtgt	2460		
gaacctggga	ggcagagct	gaggttaagcc	gagatcagcc	cactgcactc	cagcctgggc	2520		
gacagagcga	tactctgtct	aaaaaaaaaa	aaaaaaaaaa	aaagaaaataa	agtgaacaat	2580		
aaatgtatg	tggctgaatc	atccccaaac	aatcccccca	ccccagttca	cgaaaaaatt	2640		
ctccccaaaaaa	accagtcctc	ggtgccaaaa	agttggggaa	ccgctaatact	aaataatcta	2700		
atcttcattc	aatgctaaaaa	aatgaataaa	ctttttttta	aatacacccgt	ctcactttgt	2760		
tgcccaggct	ggagtaacgg	ggcatgatca	cagtcactg	tagcctcaat	caccaggcc	2820		
ccagcgtatcc	ttcccacctaa	acttccctgag	tagctggac	tacaggcactg	caccaccatg	2880		
cccagctaat	ttttaaattt	tttataagaga	tgggggtctc	accatgttc	ccagactgg	2940		
ctcaaaacccct	gggctcaagt	gatcctccct	caaactcctg	gactcaagtg	atccctcttc	3000		
cttggctcc	caaagtgc	ggattacaag	catgagccac	tgtacccagc	tggtataaaca	3060		
ttttaagtctg	cactacagtc	atggacaatc	aggctttca	acatgcagta	tggacagtga	3120		
gtcccagggt	ctgttttcc	atactgaaat	acatgtgata	ctaaggagaa	aggtgctcgc	3180		
aaggatattt	aaaatgaaga	atattaaaa	tgaggaaaaaa	actgttttt	catgactttg	3240		
ataaggctga	taaagaccat	ttctgtgatc	tcaggtgatt	cactcaagta	gtatatttca	3300		
gtaatcatta	tctggaaacag	cctgaatctt	aaccaaaata	ccatgatttt	ttaatgtctgt	3360		
tatgataacct	tgatgatatg	accaaactgc	aatgttaggca	gctaaatctc	cacgagtttg	3420		
acttccccga	gagttgacat	ttttcttcac	aaattaaaaga	aatatatttt	ttgatacatg	3480		
atggcatat	ttaaaaacta	cactgaaatg	ctgaaaaatg	atataaagaa	acattttcca	3540		
gaatcaaatg	caatcaaaaga	gtggatttag	aatctactca	ccattatcaa	ctaaatagaa	3600		
acacttggac	ttgggtgttgt	ggctcacatc	tgtaatctca	gcactttggg	aggccaaggc	3660		
aggtggattg	cttgaggcc	ggagctca	accagcctga	gcaacatagc	aaaactctgt	3720		
ctctacaaaa	aaaaaaaaaa	attaaccagg	catggtgcc	gatgtttgt	atccctcgta	3780		
ctctggaa	tgaatgatg	ggactgttg	agcccaggag	atcaagactg	cagtgagccg	3840		
tggtcatgt	ggcccacagc	ctgagtgaca	gagagagacc	ctgtctaaa	aacaaaaaca	3900		
aacaaaaaac	acttaacctt	cctgttttt	gtctttgtt	ttgtttttgt	tttgggtttga	3960		
gatggagtct	cactctgtt	cccaggctgg	agtgcagtgg	cgtgatcttg	gctcaactgca	4020		
agctctgc	cccggttcc	ccgcatttc	ctgcctcagc	ctcccgagta	gctgggacta	4080		
taggcctcg	ccaccacgccc	cggctacttt	tttgcatttt	tagtagat	gggggtttcac	4140		
cgtgttagcc	aggatggct	tgtatctctg	acctctgtat	ccacctgtct	cgccctccca	4200		
aagtgtggg	attacaggca	tgagccacccg	cacccggcca	acctttctgt	tttttagttt	4260		
gatatgttt	ttaactcagc	agctgaaaga	atgctgaaag	ttggcttcag	aaaaaaaaatt	4320		
tcactagaat	ctctcatatcc	atattaaatc	tgaatgcata	tccagattga	tcatgttagag	4380		
caaaaacact	catcatattt	cctgtatgacc	tctaattctg	tttgcgtt	tcttatttcaa	4440		
tggaaacaga	ataaggaaag	aaatgaaagg	gtctggaaa	tttgcgtt	gttatagata	4500		
ctatcaaaga	tcaccaacaa	taagatctt	cctataaaata	taaaaacaatg	ataattaatt	4560		
ttttattat	tttttctct	tcagaggatt	ttatctcaag	ataaaacata	acttctaccc	4620		
atactattga	ttccaaaggt	tagaaaaaaat	gtttttctc	atcttatactt	tcaaagaggt	4680		
cacagcaat	caaacatcta	taaaatgcct	ctgcataatt	gtcagaagct	atagtccaga	4740		
aatcattgaa	aatgttttc	cattttaaac	ttaggtgagg	tgttttagga	aacctctatg	4800		
acaacttact	ctatttatgt	ggaggttaaac	tcccagactc	tcccagggtc	tcctgtattt	4860		
atctcatttt	tttaggcttcc	taatcccttg	aagcacaatc	gaaaaagccc	ttgatctctt	4920		
ttctgcacat	atcatecgccg	aattcattcg	gttccagca	agctgacact	ccatgatata	4980		
agcggcctcg	cccttctccg	gacgcagtc	cttgcgtccgg	ttagcttagga	tgaggggttt	5040		
gctggcttc	agtgcaggt	tctgcgggtt	cccaagccgc	accagggtgc	ctcacaggct	5100		

63/122

ggatgtcacc attgcacact gagctcctgg caggctgtac caatttttta attatttaat	5160
atttattttt aaaattatgg tgaatatttt ggtattctgc tctaaaatag gcccataaat	5220
gcacagcaga tatcttgg aacccacagc ttccactgg aagaactaag tattttctt	5280
ttaaagatgc tactaagtct ctgaaaagtc cagatcctct acctcttcc atcccaaact	5340
aagacttggaa atttatgaga gatctagcta acagaaaatcc cagacacatc attgggttctt	5400
cccagagtgc agtcttccta aagaggctca gccctaagca ggcccctgca ccaggagggt	5460
gggtctgaga cccacatgc acttcccaag gtgcgtctc cagagaggca ctgaaacagc	5520
tgagcacaag cctgcaagcc tggagaactc tcacagttag aacgggggg gcccagtggg	5580
actaacataaa agaaaaaagg gaacacagag aaatggatgg caccacaaac cagaaagcc	5640
ttcatggcca atgaaagcat cagtgacggg gccagaaccc tcatccccaa agactcttca	5700
ctgccttag tgaaaaacaa tggctagaga gtgaagttat gatcatgtat agagaggtaa	5760
agttacattt ttatattctg actctgtctt gtggaaattc cctatctgtc agactaaaag	5820
tttcagacac cctgttcataa tatccattt gtgtctagag actttaaatg aacagaacgc	5880
acattgtcag gatgactatt aaaaaaaaaat caaaagacag caagtattgg tgaggatgt	5940
gagaaacttgg aacttttgtt cactgtttat gagaatgtaa aatggagcag ctgtctgtgg	6000
aaagagtatg caggttcctc aaagagtaaa accaagatgt ggaaacaact aaatgcccatt	6060
cagtggatga aggggttagac aatattgtgtt atatacatac catggagtagc tattcagcct	6120
ctaaaaaaa aaaaggaaat tctataacat gcaacagcat ggtgaatct tgaggacatt	6180
ttgctaatga aataaggcag tcatagaaag acaaataactg cacgactcca ttatatgag	6240
atacaaaaaa tagacaaaatt catagaatca aagagtacaa tggaggttac ctggagctgc	6300
agggcgaaaa acgaggagtt actaatcaac gaacataacg ttgcagttaa gtaagatgaa	6360
taagctctca agatcagctg tacaacactg taccttagatg caacaataat gtattgtaca	6420
cttaaaaattt tgtaaagggtt agattaacaa atgttagata tccacaaatg tggttaagt	6480
ttcttaccac agtaaaaataaa aaaaagataa tcaagcccgag gagttcgaga ctgcctggg	6540
taacatggtg aaaccctgtc tctacagaaa atacaaaaat tagccagctg tgagggtgca	6600
ctcctaggga ggctgagggtg ggaggcttgc ttgagccag gaggctcaagg ctgcagtgag	6660
ccatgattgc accactgtac tccagccag atgacagagc aagacaccac cccccccaaa	6720
aaaagaaaaaa gaatatcaaa cattttaaaat gatcagatac gcaagaacaa caacaaaaaa	6780
gagatgaaca gaggatcgac cctcatctg tggattctt ggtctaactg aaaaacagac	6840
attgagagac aaacaatgac agtgtatgtca tcaacagcaat tacacaggtt tccctgggg	6900
actgcagaag aaaggaggaa tgcctactt tcaaaaaata gagaagcgt caaacagttg	6960
gtgaaagcct tccaaaacta gagagaactg cacacaccaa atcacaaaaaa gaagaaaagc	7020
cgtggagat tctgggaccc accggctt ttgtatggctt gaacaccctg ctgcaggaga	7080
gacaggagct gaaagacatg gtgggatgaa acctcaaaaca gcttgcctg cattgctta	7140
gatgacttggg cttgatatac tctagtaat gggacaattt caatcaaaa agaaagatgc	7200
tcaaatttac attttagaat gatTTTTT ggcagttatgg ggaatagatt aaaagaggt	7260
gaagctggag gcaagaaaact tgtaaaggagg caactgaaac agtctatgt ataaataata	7320
aactgacaga gtgactgaaa aaatcagaac aggctgaatc aacagatacc tagatgaaaa	7380
taacaggact tgcataccatg ttgtatctt gggggatgggg agttgttcc ttgtttccc	7440
taacacttggg aatacggaa gtttgcctgt tgattttggg atataacttggt gtgtagccaa	7500
tcactgacaa ccatttagca gcttaaaaaca caaaggctta tctccagtt tctgtgggg	7560
aggaatctaa gataggctta gctggctgt tctggctcaag agtttctcaaa gaggttgcaa	7620
tcaagatgtc agctgggggtt gcatcatctg aaggctcaac tggggccgga gggccactt	7680
ccaaagggtt cactcacctg cctgacaagg cagtgtctgtt tggggcagg agatctcaat	7740
tcattggccaa gtggccttctt ctatacgatt gctggaaatcat cctccccatc tgccagttgg	7800
cttcctctcag catgagtgtat ctgagagaga gagcaaggag gaagccacag tggctttcc	7860
actctactc ctaacactat ggacctactc ctaacactct cacttcgtcc ttattccatt	7920
agttagaaaag ggaactaagc tccacccctt gaaataagaa gtgtcaaaa atttgtggat	7980
atattttaaa atcatcacatc tggaaatggggt tcaattaaatg ctgaacttga	8040
aatggcttagt accattcaat gtcaccaagg caatgaacat acccatagat ggtcatgact	8100
tttagaagaaa tagaggaaga tcacagaattt aaggaggaat tggaaatggtaa aagaagtgg	8160
gtcagattcc ccctgaaaag tgagccatga aaggaactt aactatggg ttagaggtca	8220
gagtaggaaa ttccgggtgg attttttttt aagaaaaaggaa accatataag catgttttga	8280
ggttagagggaa gaataaaatca gttagacaggg agaggtaaa aacataatg ataggggata	8340
gttgacaaaag gtcttggcag aatcccttac ccattgtactt gggggccaaag gagggacact	8400
tctttgttttgg agggataagg aaaataagaa agaattgggtt ctatgttgc tggctctgtc	8460
tcttagggcaa acgcataatg aacaaactgt gtgtgttagg aatataatg tggctccaca	8520
ttgagattct cacctcaat ccattttgtt gttacctgtt ctttccatc ttctttttt	8580
gctacatgca gactgtctttt ttgttccctt ggcctgttcc aggtttcagc attctggcat	8640
atctgttacc ctgtttccaa acctctctatc agtccatgtc ctttccctgg atagtgtttt	8700

atggggccac	gtatctaaga	agtgatgcct	tcagtttaggc	ctgagaacct	cctctatgg	8760
aatctccatc	agtgaccctg	acagacttgg	tatcttggag	atgtcactgc	tcccagcctg	8820
tggctctagga	gaatctcagc	ctgggcctct	agtagatgg	ataaggcgtt	aaggatctt	8880
tgaaccagag	tctgtcatat	tcctcaatgt	gggacagata	aaacatgtgt	agtgtctgg	8940
tttctgagct	agaactctgg	tttttggtct	agattcttt	atgtatgacc	tttcagagg	9000
attaaaaattt	gttctaatac	aatgttcaat	acaatgtag	ttccctttct	gttaggacct	9060
caacaaaaca	tgaccaactg	tagatgaaca	ttaaactatg	acaattcatg	gaaatgaata	9120
cagtaatacc	tgcggcccc	ccatTTtagc	agtcaactatg	gtgacatTT	gcacaaatgg	9180
ctattaagg	gtgttttgc	aaaacccatc	catctacta	ggcacatgt	attgaaacta	9240
atgaaaataat	ggagaaactt	ctttaaaact	tttaatgaat	aaagtgtatg	agtgtataata	9300
tttttagctgc	tatttataaa	gtgtactatta	caggtaaaac	atttcttctag	ggttttttg	9360
ttgaagttgt	cacatttaat	ccttaataaac	ccactatgag	tcaggtattc	ttctctcccc	9420
tttggacagt	tggggaaatg	gggggtcagag	aggttaggta	atttgctcg	ggccacacaa	9480
cctgcatgt	gaaaatctga	gattgtaca	ggaacgtatc	aaactctgaa	gtccatgttt	9540
ctatTTccc	atgtgcctt	tctaataaaa	ggtaataat	gtactgtat	gtgtccccca	9600
aagtggatca	ctttcacccc	accctacttg	atttctcca	ttttactaaat	cacatctgt	9660
caacttattt	attgtctgatc	tcccccaacta	gattataaaac	tcaataaaaag	caagatcctt	9720
gtctgctgaa	tatcagtacc	taaaacgtgt	tctagcacag	agcaagtaat	taatatttgt	9780
tgaatgaaca	aataaaaggaa	aaaaatttcaa	aggaagaaaa	agcccttaaa	catatgttta	9840
cctaaacata	cattttaaaa	gaaagcatat	aacaaatcc	ggacagaaat	taatattgtat	9900
tttttaaaga	aataaaccaa	tgctagctgg	gcacagtggc	tcacactgt	aatcttagca	9960
ctctgggagg	ccgaggcagg	catgtaccc	gaggtcaaga	gttcaagacc	agcctggcca	10020
acatggtcaa	acctgtctct	actaaaaata	cagaatttat	ccaggcatgg	ttgcaggatcc	10080
ctgtAACCC	agctactctg	gaggctgagt	caggagaatt	ggttgaaccc	aggaggcaga	10140
gggtgcagt	ggccaaaggat	gcacactgc	acttcaggct	gagtaacaaa	gcaagactct	10200
gtctgaagga	gaagggaaaga	aagaaggaaa	gaagggaaa	agggaaagaa	gaaagaaggaa	10260
aagaaaagaa	gaaaagaaaga	aagaaaagaa	gaaagaaa	agggaaagaa	gaaagaaaaga	10320
aagaaaagaa	aagaaaagaaa	gaaagaaa	accaagtgc	tatttgggac	ctactatgt	10380
atgttttcc	atgcacgcta	ttttcagtaa	agcagttagc	aaacttgc	gatcataaca	10440
acaatataat	gcttctataa	ctctaaaatt	gtgtttaag	agtttctct	ttaccagtc	10500
atgtatgcat	tagtttctt	agagttacta	gtactttt	cccttggagaa	tatccacagc	10560
cagtttttt	accAAAGGA	ggatgtttac	taatctgaa	tttcaaatg	ttgctcaag	10620
ttggggccagt	tcatgttaat	atactccaga	acaaaaacc	ttctactgtc	ctctgacaat	10680
tttacctgaa	aattcatttt	ccacattacc	aaggagccag	ggttaggagaa	tatagaaaaga	10740
ccacccaaga	atccctactt	ctttcagcaa	aatcaattca	agtaggtaa	ctaaacacat	10800
gccctaaca	tgaatagcag	attgtgtc	gaagaatgt	ctacaaacat	ttactgtgaa	10860
ggaactactg	aaatatttca	ataagacttc	tctccaaaat	gattttattt	aatttgattt	10920
ttaaaaaataa	ttttttagct	aaattttaaa	aggtttgata	ttgggtacatg	aatagacaaa	10980
cagacatgg	ctagaccaag	aatttaggtt	aaacatatac	aggaattttaa	tatacgataa	11040
atcttagtatt	ccaaaggaaac	caacaaatgg	tgttcagaca	gcaggatagg	catcaggaaa	11100
aacacatgg	ggcacccctac	cttactcota	acaccaggag	taactgtgaa	agcaccacaa	11160
atttattttt	tttattttata	gtttttaggt	ctaggttacg	tgtgcacaaac	atgcaggttt	11220
attacatagg	tatacatgt	ccatgttgg	gaggacacc	aaatattttaa	aagaaaaaaa	11280
ttggccaggg	gccccggctc	acacctgtaa	tcccagact	ttggggaggcc	aagggtggca	11340
gatcacctga	ggtcgggagt	tcgagaccag	cctgagcaac	atgggagaaac	cccatcttcta	11400
ctaaaaataac	aaaatttagcc	aggcatggt	gcacatgcct	gtatcccag	ctacttggg	11460
ggctgaggca	gggaaatagc	ttaatctgg	gaggcacagg	ttgcgggtg	ctgagatatt	11520
gcactccacg	ctggcacaaca	agagccaaaac	tctcaatcaa	aaaaatttaat	aaataataaa	11580
aaataaaagaa	agaaaagaaa	aaaatgaaaa	tagtataatt	agcagaagaa	aacaccgtag	11640
aatcctcgga	ctcttaggt	ggggaaatggc	tataatataa	aaaccctgaa	gttataaaaag	11700
agaaaaatacc	ctacatacaa	accaaatttt	tctacatgcc	ttttacatag	cacaaacaca	11760
gctaaataat	catatgtgaa	tgaatggaa	aaacaaatct	tgactcatat	ccagacagag	11820
tttaatttcc	tacacataaa	gagttacat	aaacacccaa	aaaaaaacc	accactaacc	11880
caaaataaaa	atgtgacagg	taatgtacag	gtagttcaca	gagaatacaa	atggcttcc	11940
ggcacataag	atgtcagac	tgacttttac	tttatttttt	tttgagagac	agggtctcac	12000
gatgttgc	aggtaggtct	caaactcc	ggctcaaatg	atagtaccag	gactacaggt	12060
ttggcccccac	gcacccgtgt	cctcaaccac	ctgttataac	agggaaatgc	aaataaaaact	12120
ttcaaatcta	ttttacat	tagatggca	aaaatttgg	aaacttcaaa	catcatcat	12180
ttgttgagaa	tgtgaggaga	ttggcactt	catttttgc	tgatgtatcata	ttatatactgaa	12240
ttggcttctat	ggaaagcaat	ctggcagcgt	ctatcaaata	tacaagtgc	tatatacttt	12300

gacaaagcaa ttccactcta	ggaatgtt	ctatatggtt	gtgcttcctg	gggctggaa	12360
ctgggagcta agggacaggg	gcagaagata	atcttctttt	ccctccttcc	ccgttaaaca	12420
tgttgattt tatatactgt	aatatattat	ttttcacaaa	agataatttt	taagcgatat	12480
gtctggaat tttttttttt	cttttcttag	acagggtctc	actctgtcat	ccaggctgga	12540
atgcctatgt atgatctca	ctgactgcag	cctcgacctc	ctgggttcaa	gcaatcctcc	12600
cacccatgcc tcctgagtag	ctgggactac	aggcacgtgc	catcatgcta	attttgtat	12660
atacagggtc tcactatgtt	gcccaggcta	atgtcaact	cctaggctca	agaatccac	12720
ccacctcagg ctccaaagtg	ctgggattac	aggcgtaggc	caccgcgcct	ggccctggga	12780
attcttacaa aaaaaaaaat	atctactctc	cccttcttgc	aaagtcaaaa	cagagaagga	12840
aattcaacat ataataaagg	tagagaaggg	cctcaaccct	gagcaacaaa	cacaaggct	12900
atttctgaga caggaatttg	ctgaacaaaa	tcgagggaaag	atgacaagaa	tcaagactca	12960
cttctcggt gggcgagtg	gctcacacct	gtaatcccag	cactttggga	ggccgaggcg	13020
gacagatcac gaggtcagga	gattgagac	atactggcta	acacagtggaa	accaggcttc	13080
tactaaaaat aaaaaaaaatt	agccggcgt	ggtggcaggt	gcctgttagtc	ccagctactt	13140
gggaagctga ggccaggagaa	tggctgtac	ccaggaagcg	gagcttgtag	tgagccgaga	13200
tcacgcccact gcacactccagc	ctgggtgaca	gagcaagact	ctgtctcaa	aaaaaaaaaa	13260
aagactcatt tctctagatc	ttgagccgt	ttcaaaattta	tctcagctta	gtgagaggtt	13320
aaagcaagga atatccttc	ctgtgggccc	tgtctcttac	tgaaggaaagg	taacggatga	13380
gtcaaggaca ccaatggaga	aaagcaactaa	caccattatc	tgtgaacat	tacgtgaaga	13440
agggtaagaa gtgaagtggaa	attgtcttgc	aagtctgtga	aaggcgacat	tcattttgggg	13500
aaatggataa tagggaaatcc	ataaaagtga	ttaaaaagat	gttagaggct	gaggcggggg	13560
gaccacaggg tcaggagatc	gagaccatcc	tggctaacac	ggtgaaaccc	catctctact	13620
aaaatatacaaa aaaatttagcc	aggcggtgt	gcaggccacct	gtagtcccaa	ctactcggga	13680
gactgaggca ggagaatggc	atgaaccttg	gagacggagc	ttgcgttag	ccgagatcac	13740
gccactgcac tccagccttg	gtgacagatg	gagactccat	ctcaaaaaaa	aaagtttagat	13800
acgagagata aagatccaaac	agacacacaa	ctgctaattc	tgaacagaac	aaaacaaaatg	13860
gcacaggaaa agaaaaattta	agatataaca	ccggaaaact	ttcctgaaat	tgagtaactg	13920
aatctatagc ttgaaagggtt	ttagcatatg	ccaaagaaaaa	tcagtaggt	ccaaaccagca	13980
caagacacat cttagcaagc	tggtgatttc	tttttttttt	gttttttttt	gggtgaccca	14040
taatcgggaa aaaggcagac	catctcgat	tttctccaga	acactgtgt	ctgaagacaa	14100
aagaatgtcg cttactgtgc	cagaagggag	agaaaagtgc	ccaaacacatc	tttaccaagt	14160
tagaatgtca cgcattattt	aaaggctgc	aaagccatga	aagacatgaa	agaaacacaag	14220
catttacaac atgaaagaaac	acaaggatcc	tcataactcaa	gaatccctta	aaaaaatgtta	14280
gtcctaattcc agccccactgc	aatgttaatg	tacttaatgt	gtctttaat	gggaacttca	14340
tagttcaaa tcagtctgtt	cccatatacc	aacatctctc	gccccgttt	cctgcaatag	14400
tcagcacctt tccctcttcc	cagtcttgc	ccctggagtc	tgctctcagc	atagcagagt	14460
gaccacatca acacccaagt	cagaggccctc	cagtgcgcac	tggtctacaa	agcccttccc	14520
accccccacc ccaacgtgccc	tccggatctc	tgtgacgtgt	cttctgcata	ccctagcagc	14580
cctgcctcc tcactgtccccc	tcgttgcatac	cgagaaggcg	actccttgc	tcttggctct	14640
ggccgcctcc tccacatgc	gtgagttac	tcctttaactt	actcttagtc	attgtctaaa	14700
tgtcagcatc tcaatggggc	cctccctgac	tacccttattt	aaattctaca	tactccccctt	14760
gaccatgg acctctacta	cccttattcca	cttttattct	tacaattttag	cacttgttct	14820
cttcttaacgt attctaagac	ttactcttatt	attacattgt	ttgcccacccc	ctctagtaca	14880
taaactcccg agggggcagg	attttctgtt	atttttcat	tttttttttt	cttaggacata	14940
gaacaggcga tagttcagat	tattcaatgt	tataatgaa	tgaactagca	gttagtaccag	15000
ttccagttag gcacagaattt	aaatcttaat	agaattaaat	ctcatggct	gggtaacta	15060
tggatagaaa attagatata	atttttaagaa	gccttagaaag	aaaaaattaa	taatgtaaaaa	15120
ataatattaa ttgtataaa	ataaaaaaaa	ctctgcccagg	cactgtggct	caaattctgc	15180
atccccagctc ctcaggaggg	tgagggtggaa	ggatcacttg	agaccagat	tcaagactca	15240
gccttagggca cacggcaaga	aactgtctc	aaaaaattta	aaatcttaat	tttttttttt	15300
gaattctaa agcgtcacaa	aaactggaga	ttaagggtaca	ggaagtgtga	agtaatatta	15360
ctatgctaat ggtttttttt	tttttttagaa	aggtataacc	aaaagatttc	tttctcaagt	15420
cgataaaactg agaaaagataa	gcataatcttc	caattaaacag	agggggagga	aaagccagat	15480
acaacaaaat aagatataaa	tttagttcca	gttgaaaaca	agatgttag	ttattttgc	15540
tcacactcacc tggaccttc	cccaggccaa	aaaaacactac	tgttaaaacag	ggtagaaaaag	15600
catcatctca gataaagcg	aaaaaactgc	cacagtctca	aaccacaaac	tataagcaca	15660
cacctggcca accctgccaa	gtctgggctc	agtaggagga	acgtgtgt	agcttaggatg	15720
taccaactta gacattctgt	gggatacaga	tgtccctgg	agggtcacac	catctcaaag	15780
gcacctgtaa tgcccactga	ttacagccac	catatgtga	agagaaaactc	aggcactta	15840
gagagtataa caagaaccc	atgtcatctg	agatgaggaa	tcctcagcccc	tgcaattaa	15900

ccaactcttt	agaacaactg	gcaaaacata	aatatccaca	acttttgcgtt	cagtaattcc	15960
actcttagat	atcaatccaa	agtacatgag	acagcagata	cacacacaaa	atggatatta	16020
ctgcacatt	gttataata	gcaaaaaaca	agaaataatc	catatgtctc	aataggatac	16080
tgggatcatg	agggtatgt	cccatcattc	aaccatcaaa	aagagtata	tggatgtcca	16140
cagatggaca	taaaaagctg	tgtgttacgt	gaaaacaaac	tcaagcagca	gcaggatggg	16200
cttatgatag	tcagtgatgag	ctaatttctg	gaaaaaaaaa	tctagttgt	gcacagaaaa	16260
catctgaaag	aacagaaaaa	aaactatcg	cagaatattt	agatgttta	ctaagttgt	16320
tatctatact	gcttgttaatt	tttacccaa	gcaagaat	ctttttggaa	aaagaaaaatt	16380
cagggaaataa	acatcttt	taaacttcat	gtttaaacaa	atggtgatgg	aataaaaagag	16440
ttcttattca	tcataaaacac	acacagcaca	catgcacgca	tgtgcgttag	cacacccttt	16500
acttgataaa	taccatgtt	aatattttg	tctttccctt	taggttctat	cccttcactc	16560
aaaatgcgg	tataaataaa	tgatctttt	atgtgccttc	tgccttaaaa	cactttaata	16620
taatcttaca	gtccccattt	cattatagtc	tcaaagctag	actcggctg	aaactaccct	16680
ttcatttgg	acccttattt	aaatggcaca	tacagctct	tcaaataaaa	acaaacccta	16740
ggacctgaca	ctaggcttcc	tttggctca	ctcataatgg	ccaagttctg	tgttataat	16800
acatctttt	tcattttatt	gctacatatac	caagggtttt	atatgtttt	tttattat	16860
cttaattcaa	aacacatca	cgctttttc	cagatggaaa	taaggaaaag	aaattgagca	16920
actgactgac	ttaaagggtc	taaaactata	tagtgaga	gtcagaaaa	gaagaaacac	16980
acatctccca	agtagaggt	gaaaaccagt	accattcacc	tccagggtga	gtatataaca	17040
gattacaaag	tcacccctc	taaatgttca	aactgaatcc	cataccata	cttaccact	17100
acctcgtaag	aacagcctca	gatctgtt	tagcctttt	tttagcatgc	tgaagccat	17160
aaaatgcctt	ccattcagca	agagaaacaa	gttctgaaac	actgaataat	ctgcccaggg	17220
cctatgaa	tttccactgt	gagaaatgtt	ttccactgt	tggagaagat	ccttactctt	17280
ctccacacag	gcagaaacatt	agaaaaattt	tttgattcta	tgatgcacag	cttagggatc	17340
tgttagcac	aatttaagt	caaatagtt	ttaatccctc	ctctgttcca	gaaacagtgc	17400
taaatactgt	gaatataaaa	attgaaaaga	tactctctgt	gtcccaaga	aagttagcc	17460
gatagaggag	acacaggcac	acaaatact	gtcacatgaa	gtcttaccc	cctaacttca	17520
aacgaggggcc	taagtccacca	agaatacagt	agcagttgt	actacgata	actactataa	17580
ttcaatactt	tatcttccct	tagaaaactt	tttcccttgg	gaaattattt	tgatcttctt	17640
aataccattc	tttactaaaa	ggaagcaggg	ctccttgggg	aaatagctg	ttcttaggt	17700
ggactatgaa	atgaaaatgg	tgagtcgtt	acatccat	ttgcccagaa	atcaaggaac	17760
tgcccaaaga	ttaacagagt	catgtttaat	ggacctaaga	gtgaacc	aggagctcac	17820
tttggcccg	gttggacaaat	ttcaagaaaa	acatgcacat	aatgaattat	aaaacatgaa	17880
ttaaaatata	tatttgtact	aaaaagagaa	caaaggatgt	tggcttggaa	taaagcttctt	17940
cttcatggaa	gaataccagc	taataatgt	aaagggaaatg	agagaattag	aaaaatttac	18000
atttgtaaa	ccttaatata	ttcacctaga	catgtttaaa	ccactgagta	aaaggctgct	18060
tgggaagagg	atgctcacat	gatctcagag	tttccacacca	cagataattt	attagataca	18120
ggaaggaaga	tgtgtcaag	cttcctgtga	ccccccagcca	ggccccacaa	cactatgtgc	18180
ctccctgtca	tgtggggatc	acacagatc	gcccacacag	cttctcgcca	aaactgtttt	18240
aagctaatca	caagggaga	actggacagc	ttctgaccat	gagacgctcc	accagacaac	18300
ttgcttggcc	tctccaaaga	aacttgcctt	gcctctccaa	agaaaactca	gttcttattt	18360
aaaacaaaac	taatttattt	aaaacaaac	aaaagcaagt	tgtggactt	agtcctcagg	18420
acagagcaga	cataatttc	cctgttctt	ccgttaagt	gtataaaaa	ccctcaacac	18480
tagatataaa	acaatataaa	aaagggttctt	gaagggggaa	aggaggcaga	ctatccaggt	18540
gccttgggg	ccacagaaca	accccaatgt	gggttctact	ggcttcttt	ttgcttattt	18600
atctcgact	tggagcttca	gcagcaggca	acttcaaaac	accaaggggc	acagattgaa	18660
aagccccaaag	aaaagccttc	cctctttagc	caaaggacca	ggaaggagac	agtctaatga	18720
gatggacacat	atttagacag	taacttgcctt	tttccatggca	ataactgac	aggggccta	18780
gacttccagt	tttgcgtt	cgtagca	tatccaaac	ccccccaaag	gtcgatgt	18840
gacttgcgt	tttatccctt	catggcgtt	gttggatggcc	catccctc	ccggccagcag	18900
tgtcagggg	acctggactt	ccactccac	ccaggatgt	tgaggccctc	cctgctgggg	18960
tcatgtcaga	ggaggcctag	tggagattca	gtgacttaac	ctttttccat	agataatgt	19020
gccacccctt	ctcccttctt	ccccatgtgt	acagtggaaag	cactgtggca	acgactggc	19080
actcttaccc	ctcccttagca	ggggatgtatc	agggggccca	atggggaaac	cagaataaccc	19140
acaacccccc	agcagcaaca	ggggcccccc	acccctattgg	gtgtcaatgg	aacggagcg	19200
gaaaggccttgg	atatttacc	ccatcttagaa	gttacaagct	gtatcccccc	tttcttctact	19260
acaatgggtt	tcaaaacagg	tttaataaag	gtcttagatgt	tgataacgt	atacccaaat	19320
cgttgaagtt	ttcatttgagg	atcatttata	ccaagatgt	ggaagatccc	aaactgaaag	19380
agagaaaaaa	caatttgacag	acactagcac	taagagagca	cagatattag	aactacctga	19440
aaggatgtt	aaggacat	cataaggcctc	aacaggctgg	gcgcgggtggc	tcacggctgt	19500

aaccccagca	ctttgggagg	ccgaggcagg	tggatcacaa	gatcaggaga	tcgagaccat	19560
cctggtaac	acgggtaaaac	cccgtctcta	ctaaaaataac	aaaaaaaaat	agaaggcat	19620
ggtgggggc	acctgttagtc	ccagctactc	gggagcctga	ggcaggaga	tggcatgaac	19680
ctggaaagag	gagcagttag	ccgagatcgc	accacccgcac	tccagcctgg	gcaacagagc	19740
aagacctgt	ccccaaaaaa	aaaaaaaaaa	aaaaaaaaaagc	ctcaacaac	aactacaaac	19800
gtgcttggaa	caaataaaaaa	aaaaatcttg	gcaaagaaat	aaaagatata	tatTTTggcc	19860
aggtgcagt	gctcacagcc	tgtatccct	gcactttggg	aggctgaggc	aggcgatca	19920
cctgaggctca	ggagtttggag	accagctga	ccaacatggaa	gaaaccctgt	ctctactaaa	19980
aatacaaaaat	tagccatgtc	tggtggcaca	tgcctgtata	cctagctact	caggaggccg	20040
aggcaggaga	atcgcttggaa	ctcaggaggt	ggaggttgcg	gtgagccgag	atcccgccat	20100
tgcacattgc	actccagcc	gggcaacaag	agcaaaactc	catctcaaaa	aatatagatac	20160
atattttaat	ggaaattttt	gaattttaaa	atacagtaac	caaattgtat	ggaagagcaa	20220
catagaatgg	ggggggcaga	caaaataatc	agtgaacttc	aacagaaaaat	aatagaaaatt	20280
acccaatatg	agaacagaaa	agaaaaataga	ctggccaaaaa	aataaagaag	aaaaaaagagg	20340
agcagcagga	gaaatgtatgg	aaaaagagaa	agaaggaaag	gaagggaaagg	aggaggggaa	20400
ggagtgaggg	agaaagtctc	aaagacctct	gagactaaaa	aaaagatct	aacacttgc	20460
atcagggtcc	aggaaagaga	caaagatggc	acagctggaa	acgtattcaa	aaaataatag	20520
ctgaaaactt	cccaaatttg	gcaagagaca	taaacctata	gattcgaaat	gctgaacccc	20580
aaataaaaaag	cccaataaaaa	tccacacaa	aatacatcat	agtcaactt	ctgaaaagac	20640
gaaaagagaa	aacgtcttgc	aagcagttag	tgaaacaaca	cttcatgtat	aaggaaaaaa	20700
caattcaagt	aacagattt	ttacagaaat	taaggaagcc	agaagggaaat	gacacaatgg	20760
tttcaagt	ctgaaagaaa	agaagtgtca	acacaaaat	ctagattcag	taaaaatatc	20820
cttcaagaat	caatgggaa	tcaagacat	ctcagataaa	gcaaaaataag	agaatatgtt	20880
gccagcagat	ctccccctaa	ggaatggca	aaggaagatc	atgcacacaga	ccaaaaaaatg	20940
atgaaaaaag	gatccagaaa	acatcaagaa	gaaagaaaata	acatagtaag	caaaaataca	21000
tgtatttaca	ataaaaattt	tatctcttct	taagacttct	aaattatatt	gatggttgaa	21060
gcaaaaatta	taaccctgtc	tgaagtgtct	ctactataat	tatgcagaga	attataaattg	21120
gggaaagtat	aggtttctat	acctcatttg	atggggaaaa	tgacaaact	gtgaaaagtt	21180
acatacacac	acacacgtaa	gttatataaa	atatatgtt	gttatatgt	gttatataat	21240
atatacatat	ataatgtat	acagcaacca	ctaacaacac	tataaaaaaa	gataataacc	21300
aaaaacaatt	tagataaaatt	gaaatggaaat	tctaaaaat	attcaataac	tctacaggaa	21360
gacaagacaa	aaagaaaaaa	aaagaggagg	acaaactaaa	ttttttaaaa	acataaaaaaa	21420
aatggtagac	ttaaggccct	acttataat	attatacata	atgtaaatga	tctaattata	21480
tcaataaaaa	gacagagata	gcagagttaa	ttaaaaaaca	tagctataag	aaacctgttt	21540
tggctgaggt	gcagtgcact	acacttggaa	tcccagcact	tcgggaggcc	aaggcgggtg	21600
gatcacctga	ggtcaggagt	tccagaccag	cctggacaaac	atggtaatac	cccatctcta	21660
ctaaaaatac	aaaaaaat	gccaggcat	gtggcacac	cctgtatgc	caactactca	21720
ggaggctcg	acacaaagaa	tgcttggaa	ccggcagcag	aggttagcgt	ggccaagat	21780
tgcgcactc	cagcgtgaa	gacagatgt	gactccac	cagtggaaaa	acaaaaaaaga	21840
aacctgttt	aaatatacc	acatatgtt	gttggaaat	aaagaataaa	atatatcatg	21900
aaaacattaa	tcaaaagaaa	ggagtggtca	tattaataac	ataaaataga	cttcagagaa	21960
aagaaaattt	caagagacag	gaataaaaagg	atcaagaaaa	gatcgtgaaa	gaaaaggcagg	22020
caaataatc	attctgttg	gagatcaac	accctcttct	aaaaactgt	agaacaacta	22080
gacaaaaaaa	tcagcatgta	gttggaaaga	acttaacac	actgaacac	aggatctaat	22140
agacattttac	ggaaacactt	acccaaacaaat	agcaaaataaa	acattcttt	caagtattca	22200
ctgaacat	ccttagaccc	taccctgggc	cataaaaacaa	agctcaact	tgattgccga	22260
aggcttggat	ggacagtgg	agagctgcat	ggggagggag	aagggtgac	ttaaagagtg	22320
taggatttct	ttttgggat	atggaaaatgt	tccaaaaattt	atgtggtg	tgttggcgca	22380
actctacaaa	tataaaaaaaag	gccattgtat	tgtacgttt	aagtgggtg	aacatatgtt	22440
atgtggat	tatctaacgc	ttttttaaaa	ttaaacaacat	ttcaaaaat	agaagtctata	22500
cagagtgtgc	tctactggaa	tcaaaactaga	aagaggttaac	tggaggataa	cgagaaaaagc	22560
ctccaaatac	ttgaaaaact	gacagcacat	ttctaaaatc	atccgtgggt	caaagatatt	22620
catttctgtat	attcttttt	attgtttat	gtatTTTaa	aaatttttta	agggaaataaa	22680
actgactaaa	aatgaatatg	gctgggtgc	gtggctcag	cctgtatcc	cagcacttt	22740
ggaggccgag	gctgggtggat	cacaagatca	ggagttcag	accagcctgg	ccaagatgg	22800
gaaacccctgt	ctcaactaaa	aaactacaaa	aagttagccaa	gcmcagtggc	ggggcctgt	22860
ggtcccagct	acttgggagg	ctgaggttag	agaatcgctt	gaacacag	aggagaggtt	22920
gcagtgcagcc	aagattgtgc	cactgcacgc	cagcctgggc	gacagagact	gctcaaaaaa	22980
aaaaaaaaaa	aaaaagaata	tcaaaatttg	tgggacata	ttaaagaaat	gctgagaggg	23040
aaatttataa	cactaaatgt	ttacatttga	aaagagaaaa	agtttcaat	caatagtctc	23100

68/122

cactccatc	tcaagaacac	agaagatgaa	gagcaaaaata	aacccaaggc	aagcaaaaga	23160
aagaaaatat	aaaataaat	cagtaaaatt	gaaaacagaa	acaacataaa	gaaaatcagt	23220
gaaacaaaagt	actgattctt	cgaaagatta	ataaaaattga	caaacccta	gcaaggctaa	23280
caaacaaaaa	agaaagaaga	cacggattac	cagtatttag	aatgaaagca	taatttagaaa	23340
caactctaca	cattataaaat	ttgacaatgt	agatgaaatg	gactaattac	tgaaaaaaaca	23400
caaattacca	caactcaccc	aatatgaaat	agataattgg	gatagccgt	taactactga	23460
gaaaattgaa	tttgtaattt	taacacttct	aaaacagaaa	cattaaacct	aatattttat	23520
aaatattaga	taaggttaat	atacccttc	ttaacaaata	aaaacgacaa	attattttgc	23580
agctaaagag	atgtatgtac	tgtaaaaat	atcttcagaa	aatagaact	ttgtttgaag	23640
aataaggatt	taaaaaatgt	ttttaactct	caagaagcaa	atatctgggc	ccagatggtt	23700
tcactgaaga	attctaccaa	atgtttatg	aagaattacc	accaactcta	catagcatct	23760
ttgagaaaaac	tgaagagaag	ggaacatctc	ccagttcatt	ttatgaatg	gggttactc	23820
tgatactaga	actgtataaa	gacagctact	cttgacacac	tgcctatggg	tagctctgct	23880
ctgcaggaac	agttagaaaa	aaaaaaaaaa	gaagactgg	acaaggcag	tataaaaaaa	23940
gaaaactggg	ccaggtgcag	tggctcacac	ctgtatctc	agcactttgg	gaggctgacg	24000
ctggggatc	acctgagggtc	aggagttga	gactagcctg	gccaacatgg	taaacccttg	24060
tctctactaa	aatacaaaaa	ttagccaggc	agggtgggtg	ggaaaataaa	aaggaaaaaa	24120
aaacaaaaaa	aaactgcaga	ccaatatccc	tcatgagat	agacacaaaa	ctccttaaac	24180
tccttaacaa	aatattagca	agtagaaagca	atataaaaa	ataattataac	accatgatca	24240
atggggactt	attccagaaa	cgcaagttc	gttcaacatt	tgaaaacaag	gtaacccact	24300
atatgaacgt	actaaagagg	aaaactacat	aatcacatca	atcaatgcag	aaaaaagcat	24360
ttgccaaaat	ccaatatcca	ttcatgatcat	tctaataaga	aaaataagaa	taaaggggaa	24420
atccctgac	ttgataaaggc	ttacaaaaaa	ctacaaaaagc	ttacagctaa	cctatactta	24480
atgggtaaaa	actaaatgtc	ttccccatcg	atccggaaaca	aagcaaggat	gttcaacttc	24540
attgctctt	ttaaacatag	ccctgaagtt	ctaacttgc	caaaacgata	agaaagggaa	24600
atgaaagacc	tgcagattgg	caaagaagaa	ataaaaactgt	tcctgtttgc	agatgacatg	24660
attgtctcat	agaaaaatgt	aagcaactag	gggttaggggg	gcagtgaggaa	cacgctggc	24720
aaaggatacc	aaatttcagt	tagggaggat	aagttcaaga	tacatttgc	acaacatgg	24780
aactactt	aatatattgt	attcttggaa	atactaaaaag	atgggtgtt	aagcggtctc	24840
acgacaaaaaa	tgtataatcat	gtgaaat	gcatacgta	attagcacaa	cgtatattac	24900
tccaaaacat	catgttgcac	atgataaata	cacacaattt	tatctgtcag	ttaaaaaaca	24960
catgattttg	gccaggcaca	gtggctcata	ccctgtatcc	cagcatttta	ggaggctgag	25020
gcfgagcagaa	aacttgggt	cgggagttt	agaccagaat	ggtcaacata	gtgaaatccc	25080
gtctccacta	ataatacaaa	aatttagcagg	atgtgggtgc	gtgcacccgt	agacccagct	25140
acttgggg	ctgaggcagc	agaattgtt	gaacaaaggga	ggcagagggtt	gcagtgagct	25200
gggtgcccact	gcattccagc	ctgggtgacag	agtgagactc	catctcaaaa	aaaataaaaat	25260
aaagcatgac	ttttcttaaa	tgcaaagcag	ccaagcgcag	tggctcatgc	ctgtatccc	25320
accacttgg	gaggccgagg	caggcagatc	acaaggtcg	gatgtttgaga	ccagcctgac	25380
caacatggta	aaacccatc	tctactaaaa	aatatataaa	ttagccaggc	atgtgtagtc	25440
tcagctactc	aggaggctg	ggcaggagaa	tcacttgcac	ccggaggcag	aggttgca	25500
gttggccac	cgcactccag	cctgggtgag	agaacgagac	tccgtctcaa	aaaaaaaaag	25560
caaataacc	taattttaaa	aacactaaaa	ctactaagt	aatttgcgt	gtcttttagga	25620
ttcaggatat	atgtgtacaa	tacaaaaat	aatttgcgt	gacaaggag	gatttttta	25680
ggtcgtatgt	tttggggctgt	aatgcacat	gatttgcct	gtcaatagct	gctgtgtcc	25740
agcctgagca	gcataatgt	accacatctc	tatttaaaaa	aaaaaaaatt	gtatcttat	25800
gtactagcaa	taagcacatg	ggtactaaaa	ttaaaaaacat	aataaaatact	gtttttatt	25860
gcctaaaaaa	aatgaaataac	ttacatataa	atctaacaaa	atgtgcggaa	cttgggtgtct	25920
gaaaactaca	aaacgcgtat	aaaagaaaat	aaagaagact	taaaatagct	gaaatataacc	25980
atgtttatag	tttggggaaac	ttaatataat	aaagatgcca	atttatccaa	aatttataca	26040
caggataaca	tttattactac	aaaaatccca	gaaaattttt	acatagat	agacaagatc	26100
atacaaaaat	gtatacggaa	atatgaaag	gaacttagat	agctaaaaac	aatttggaaa	26160
agaaaaataa	agtgggaaaga	atcgtctat	ccagtttcaa	gacttacata	gctacagtaa	26220
tcaagactgt	gatattgaca	gaggcacacg	tatagatcaa	tgcaacccaa	tagagaacta	26280
agaaagaagc	acacacaaat	atgcccataat	gatttctgac	aaagggtt	aaacacttca	26340
acggggaaag	atatgtctt	cattaaaggg	tgttagatca	ttgcacatct	ataggcaaaa	26400
agatgaacat	gaacctcaca	ccctacagaa	aaattaactc	aaaatgactc	aaggactaaa	26460
cataagatat	acatctataa	aacattttaga	aaaaggccac	gcacgggtgc	tcacgctcg	26520
aatcccagca	ctttggggagg	ccaaggcagg	tggatcacct	aaggtcaggaa	gtttgagacc	26580
agccggatca	acatggagaa	gccccatctc	tactaaaaat	acaaaattag	ctggacgtgg	26640
tggcacatgc	ctgtatcc	agctacttgg	gaggctgagg	catgagaatc	gttgaacccc	26700

ggggggcaga	gttgcggtg	agccaagatc	acaccattgc	actccagcct	ggcaacaag	26760
agcaaaactc	caactcaaaa	aaaaaaaaaa	aaaggaaaaa	tagaaaatct	ttgggatgta	26820
aggcgaggta	aagaattctt	acacttgatg	ccaaactaag	atctataagg	ccagtctgtgg	26880
tggctcatgc	ctgttaattcc	agcactttgg	tcaacttagat	gaaaggatata	ttgggattca	26940
ctgttatttt	cttcaactt	ttctgttaggt	ttgacatttt	tttagtaaaa	aattggggga	27000
aagacctgac	gcagtggtctc	acacctgtaa	tcccagcact	ttgggaggcc	ggggcagggtg	27060
gatcacacgg	tcaggagtt	gagaccagcc	ttggccaaatc	ggtaaaaccc	cgtctctacc	27120
aaaaataataa	aaaattagcc	gggtgtcatg	gtgcatgcct	gtaatcccag	ctactgagga	27180
ggctgaggca	ggagaatcac	ttgaacctgg	gagggtggaa	ttgcagttag	ccgagattgt	27240
gccactgcac	tccagccttg	ggtgacagag	cgagactccg	tctcaaaaaga	aaaaaaaaaa	27300
aaagaatatac	aaacgcttac	tttagaaact	atttaaagga	gccagaattt	aattgttatta	27360
gtttagag	caatttttat	gctccatggc	attgttaaat	agagacaacca	gctaacaatt	27420
agtggagttc	aacagcttt	aaatttgcata	actgttttag	aagagagccc	tatcaatatc	27480
actgtcattt	gaggctgaca	ataagcacac	ccaaagctgt	accccttga	ggagcaacat	27540
aagggttta	accctgttag	ggtgttaatg	gttggatat	ggttggttt	ccccaccgaa	27600
gtctcatgtt	gaaatttgg	ccccagtaact	ggaggtgggg	ccttatttgg	agggtctga	27660
gtcataaaaa	ttggcatatcc	ctccatgtt	gttgggtggcc	attcttgcag	aatgtgagtga	27720
gttcttactc	ttagttccca	caacaactgg	ttattaaaaaa	cagcctggc	cttccccca	27780
tctctcgctt	cctctctcac	catgtatct	cactggttcc	ccttccctt	atgcaatgag	27840
tggaaagcgc	ctgaagccct	cgccagaagc	agatagtgtat	gccatgttcc	ttgtacagcc	27900
tacaaaacca	tgagcccaat	aaacctttt	tctttataaa	ttatccagcc	tcaggttattc	27960
ctttatagca	agacaaatga	accaagacag	ggggaaatca	acttcattaa	aataatctat	28020
gcagtcacta	acaaaataag	aacaagggc	tccagaagtg	ggaaagccat	accaggagtt	28080
cctacaatac	agtatctgaa	aagtccagg	tccaaacaaa	aaatataat	atacaggccg	28140
gacatggtag	ctttagtgcgt	taatccagc	actttggat	gctgaggcgg	gcagatcacc	28200
ctaggcagg	agttcgagac	cagcctggcc	aatatggca	aacccctct	ctactaaaaaa	28260
tacaaaatt	agccaggcat	gggtgggtat	ggctgtatcac	ccagctactc	gggaggctga	28320
ggcaggaaat	cacttgaacc	caggaggcag	gggttgcagt	gagccgagat	cacgccactg	28380
aactccagcc	tggcaaccaa	agtggagactc	cacccatcaaa	aaaaaaaaaa	tatacatata	28440
tatatgtgt	tgtgtgtgt	tgcgcgcgt	tgtgtatata	cacatacaca	tatatacata	28500
tatacagaca	cacatatata	tatgaagcat	gaaaagaaaac	aaggaagtat	gaaccataact	28560
ttctgtgtt	atgataggat	ggggatcatc	gggggaaatca	gacaagggaa	actgcaagt	28620
agagcaaaaca	gttacatcgat	ttaacagaaa	aagactttgg	gataaccatt	ataaatatgt	28680
ccacagaatt	aaagaaaaggc	gtgataaaaa	aaggaaagga	aagtatcata	acaatattac	28740
tccaaataga	gaatataatc	aaaggcatag	aaattataaa	atataataca	atggaaattc	28800
cggagttgaa	aggttagaata	actaaaaattt	aaaatttact	agagaaggat	caacactata	28860
tttgaaactgg	cagaagaaaa	attagttag	acaaatataac	ttcaatagac	attattccaa	28920
tgaaaaaataa	aaagaaaaaa	gaatagaaaa	aaataaacag	aatctcagca	aaatgtggca	28980
caccattaa	cacattaaaca	tatgcatact	gaggttaccc	gaagcagatg	agaaaagagga	29040
agaaaaataa	ttcaaatatg	ggccagtaac	ttcttagatt	tttggttttaa	agcaataacc	29100
tatacaatca	agaaaactcaa	tgaattccaa	gtaggataaa	tacaaaaaga	accacaaaaca	29160
gatacaccat	gttaaaaatg	ctgttaagtca	aaaacagaga	aaatattgaa	agcagctaga	29220
ggaaaaactt	taagagaacat	tcaacttcaa	agaacatcatc	cttataaaaag	aaccacaata	29280
atagaaaacag	ttgaccccttc	atcagaaaaac	ataatgtata	acatatttga	agtgtctaaa	29340
gaaaaaaaat	aaaggatcc	atatacgaca	aaatctgtct	tcaaaaaat	acatccaaaa	29400
ggattgaaac	cagggcttt	aaagaggatt	tgtacatcca	ttttcatagc	agcattattc	29460
acaatagcca	aaaggttagaa	gcaacccaaag	ggtccatcg	caaataataa	aaatgtggta	29520
tatgtatata	caatggaaatt	tattctatgt	taaaaaaggaa	tgaatattctg	acacatgtca	29580
caacatggct	aaaccttgg	aaacactatgc	taagtggaaat	aagccagcca	caaaaaggaca	29640
aataccatata	tacttcaatt	gtatggataa	ccttagggtag	tcaaaatcg	agatagaaaag	29700
taaaacagt	gttgc当地	gctggggag	ggagtaacgt	ggagttattt	ttgaatgggt	29760
acagaattt	agtttgccaa	gataaaaaaga	gttctggaga	cagatggtgg	tgagggtgg	29820
acaacaaatac	aaataatactt	tataactact	aacagtataac	ttaaaaatga	ttaacatgg	29880
gaaaccccg	ctctactaaa	aataaaaaaa	attagctgg	gtgtggggc	gggcacccgt	29940
aatcccagct	acttggggagg	ctgaggcagc	agaaattgtctt	gaaaccagaa	ggcgagggtt	30000
gcagttagct	gagattgcgc	caccgcactc	tagcctggcc	aataagagca	aaactccgtc	30060
tcaaaaaata	aaaataaaaa	aaaattttaa	aatgatataag	caggaggc	ggcacccgt	30120
ctcacacatc	taatggccagc	actttgggg	gcccaggcag	gcatcactt	gagaccagga	30180
gtttgagacc	agccctggcc	acatggcaaa	accctgtctc	tgctaaaaat	acaaaaatta	30240
gccaggcatg	gtggcatata	tttataatcc	cagctactgg	tgagactgag	acacgagaat	30300

70/122

tgcttgaacc	caggaggcag	agattgcagt	gagtcgagat	cgcgccactg	aattccagcc	30360
tggcgacag	agcaagattc	tgtctcgaaa	aaacaaaaac	aaaaacaaaa	agcaaaacca	30420
aaaaataatt	aagcaggaaa	cgagattgt	gctgaggagg	agaaaagatgt	gcaggaccaa	30480
ggctcatgag	agcacaaaaac	ttttcaaaaa	atgtttaatg	ataaaaatgg	taaattttat	30540
atgtatctta	ccacaaaaaa	aaggctggg	gggcaggaaa	tgaagggtgaa	ataaagacat	30600
cccagagaaa	caaaagtaga	gaatttgtt	ccttagaaga	aacaccacag	gaagttcttc	30660
aggctgaaaa	caagtgaccc	cagagggtaa	tctgaattct	cacagaaaat	tgaagcata	30720
cagtaaaggt	tattctgtaa	ctatgacact	aacaatgcatt	atttttctt	ttcttctctg	30780
aaatgatttta	aaaagcaatt	gcataaaaata	ttatataat	agccttattt	tgaacactata	30840
acatataatag	aatatactt	gtatataatt	tgcaataaaac	tgcacaaaag	agagttggaa	30900
caaagctgtt	actaggctaa	agaaaattact	acagatagta	aagtaatata	acagggaaact	30960
taaaaataaa	attttaaaaa	atttttaaaat	ataattaca	acaataatata	ggttgggttt	31020
gtaatattaa	tagacataat	acaaaaatac	cacaaaagg	gaagaagaca	atagaactac	31080
ataggaataa	cattttggta	tctaatacgat	attttttat	aatatataat	tgatattctgg	31140
taagttaaaga	cacacatgtt	aaacccttaga	tactaaaaag	taactccat	aaatacagta	31200
aaaaataaaa	taaaatatttta	aaaatgtttt	tattatgtt	ctcagggtac	agtaacaaac	31260
taccacaaat	tgagtggctt	aacacaactt	aatgttattt	tctccagtt	ctggaggctta	31320
aacacacgtca	atcaaggtga	gtacaggccc	atgtccccct	tgaaggctt	agggaaagaaat	31380
cctcccttgt	ctcttccagc	ttccagttt	tctcaatgtt	cctaagtgt	ctttggcttg	31440
tagtatatac	attccatgca	accaggaaa	agaaaataat	aaagattatg	gcaaaaaaaata	31500
atgaaatcaa	aaggaaaaaa	atggaaaaaa	ataaaataaa	ccaaaagcta	gttctttgaa	31560
aagatcaacc	aagtttacaa	acctttttaac	tagactgaca	aaaaggaggt	aagactcaaa	31620
ttactagaat	cagaaataaaa	agaggggaca	ttactaatga	gggatagaa	aagaataacta	31680
cgaacaaaatg	tgtgccaaca	atttttttttt	tttagatgaa	atggacagg	tcttaggaca	31740
acatcaacta	ccaaaaattt	ctcaagaaga	aaagagaaat	ttgaatggc	tataacaagg	31800
gaagagactg	aatttgacaa	caaggaaacta	tccacaaaaga	aaatcccagg	cccaagaagat	31860
ttcactgtga	aattttttca	aacttataaa	tataaattaa	catcagttt	tcacaaaactc	31920
ctccaaaaaa	aagaacagat	ctcttatttac	aggcgatag	attttttggaa	aatcctaagg	31980
gaactactaa	gacactatga	taactgttta	acaaggttt	caaggctgca	ggatagaaaaa	32040
ccaatataca	aaaatctatt	atattttctat	acattttttt	tgaacaaacc	aaaaatgtgaa	32100
tttaaaaaat	aattcaattt	acaataacat	aaaaaagaat	aaaacactc	aaaaataaaat	32160
ttattcaagt	aagtgcaaaa	tttataactt	agaagctaca	aaacactgtt	aaaagaaatt	32220
aaaggtttac	ataaaatggaa	aactatccca	tgttcatgga	tcaaaaagact	tattactggc	32280
aatgctctcc	aaattgtatct	ataaaatttca	aaaaatccct	atcaaaatcc	cagatgaggc	32340
tgggggttgc	ggtttcatgccc	tgtaatccca	gctttttttt	aggcttgggg	aggcttgggg	32400
cctgggtcg	ggagctcgat	atcagccctga	ccaaacatgg	gaaaccctat	cttttcttaaa	32460
aataaaaaat	tagtcaggcg	ttggggcaca	tgcctataat	cccagctact	cgggaaagctg	32520
aggcaggaga	atcgcttga	cccaggaggc	agagggttgc	gtgaggccaa	atcggttccat	32580
tgcactccag	cctggggcaac	aagagcaaaa	ttccatctca	aaaaaaaat	aaaaaaaat	32640
ccagatgact	tcactgttga	aattttttttt	attttttttt	ttttttttt	ttttttttt	32700
accttgagaa	tagccaaaaac	aaacttggaa	aacacgaaaca	aaatataat	tgacttactt	32760
gccaatttgc	aatgttacga	cacagcaaca	gtaatcaaga	ctgtgtggta	ctggcaaaag	32820
acacatacat	acatacatat	caatggata	taattttttttt	ttttttttt	ttttttttt	32880
tctatgtttaa	gtgtttttt	ttttttttttt	ttttttttttt	tttttttttt	tttttttttt	32940
tcaccatgt	tcctttttttt	ttttttttttt	ttttttttttt	tttttttttt	tttttttttt	33000
cctcccaaaat	ttttttttttt	ttttttttttt	ttttttttttt	tttttttttt	tttttttttt	33060
aaaggatcaac	aaaggatccat	ttttttttttt	ttttttttttt	tttttttttt	tttttttttt	33120
aaaaaaaaaa	tgaagggttga	cccttccatca	cactaaatgt	ctgcgattat	aggcatcagc	33180
caccacatcc	agcccaaaatg	atttttttttt	ttttttttttt	tttttttttt	tttttttttt	33240
aggcttggga	taatcgatata	gtcataatgt	aaaaaaaatg	aggcttgggg	aggcttgggg	33300
ccatatgc	aaattttttttt	aaaaatgtat	ttttttttttt	tttttttttt	tttttttttt	33360
aaaactttaa	gaagggaaaaca	tacgggtaaa	tctttttttt	tttttttttt	tttttttttt	33420
cttagacatg	acacaaaaaa	catgaccaac	taagttttttt	tttttttttt	tttttttttt	33480
aaaatggaaa	acctttgtgc	ttggaaaggac	accatcaaga	aatggaaaac	aaaaatagcc	33540
aaggcaatata	taagttttttt	gaacaaatgt	ggaggatcata	tacttcttgc	tttcaaaagca	33600
acactaacc	aaacacgtat	ttttttttttt	aaaaacagac	acatagacca	atggaaacaga	33660
ataaagaacc	aaaaataaaa	tccacatatt	tatgttcaac	tgatattttt	caatgacacc	33720
ccttcaataa	atgataactt	gaaaacttgg	tatgttcaat	cagaagaata	aaacttagacc	33780
cctatctctc	accatatacg	aaaatcaact	cagactgtt	taaagacttg	aatgttacacc	33840
ccaaaactat	aaaactactt	gtagaaaaaca	taaggaaaaaa	cgcttccat	catttttttt	33900

71/122

ggcaaagatc	ttatggctaa	aacctcaaaa	acacaggcaa	caaaaacaaa	aatggaaaaaa	33960
tagacttta	ttaaactaaa	aagctcctgc	acagcaaagg	aaacaacaga	atgaaaagac	34020
aacctgtaga	atgggagaaa	atattgca	actatccatc	catcaaggg	ctatgtatcca	34080
gaacacacaa	gtgactaaa	caactcaaca	gaaaaaaaaagc	aaataatctg	gttttatata	34140
ggccaaaaga	tctgaataaa	cattctcaaa	ggaagacata	caaatgtcac	tatcattctg	34200
ccagtaccac	actgtcttga	ttacttgtt	gtgtataaaat	tttaaattt	ggaagtgtga	34260
gtcatcctac	actttgttct	tggtttca	gtttgtttt	gttattctgg	gagccttgc	34320
agtataaaaat	agcacaacaag	tatgaaaaaa	tgctcaccat	cactaatcat	cagagaaaata	34380
aaaatcaaga	ccactatgag	atatctctc	actccatgtt	aatggatc	tatcaaaaag	34440
acaaaatata	atggatctg	gcaaaagattt	ggagaaaagg	gaactccat	acactgtggg	34500
tagggatgca	aatttggata	ggccattatg	gaaaataata	ctgagggttt	tcaaaaaact	34560
gaaaatagaa	ctaccatatg	atccagcaac	cctactact	gttattttatc	caaaggaaaag	34620
aagtcatgt	actgaagaaa	tatgtca	ctcatgtt	ttgcaacact	gttcacaaca	34680
gccaaagacag	ggaataaaaatc	taaatgtca	tcaacagat	aatggataaa	gaaaatgtgg	34740
catataact	caataaaaaata	ctattcagcc	attaaagaag	aatgaaaatcc	tgtcatccca	34800
gcaacatgga	tgaaccttgg	ggacattata	ttaatgaaa	taagtaaagc	acaaaaagat	34860
aaacagtaca	tgttctca	cagacatggg	tgctaaaaag	aaaatgggt	cacagaatta	34920
gaaggggagg	cttgggaaaaa	gtaatggat	aaaattttt	agctatgtt	gaagaataag	34980
ttttagtgg	ctatagaact	gttagggcag	tatagtacc	aaaaattttat	tgtagatgtt	35040
caaaaagcta	gaagagattt	tggatgttcc	cagcacaaag	aatgataaa	tgtttgtgat	35100
gatggatata	cttaatttacc	tgatttcaatc	attacacatt	gcatacatgt	atcaaattat	35160
cactctgtac	ctcataaata	tgtataattt	ttacgtcaac	aaaaattttt	aaaaaaagga	35220
aattaagaca	acccacataa	tggaaagaaaat	aaaatatctg	caaatttat	atatctgata	35280
aatatttata	atttataata	tataaagaac	tcttacaact	caagacaac	aacaaaacaa	35340
ccccatttca	aatatggtaa	aagcttggaa	tatacactt	tctaaagact	atatacaatt	35400
ggccaaataaa	gacacgaaaa	gatgtcaac	atcactatgt	atcaggaaa	tataatcaa	35460
aaccacaatg	tagaatgttag	acaccactt	atatgcata	ggatggctg	aaaaaaaagg	35520
taataacaaa	tgttggtaag	gatgtggaaa	aatcagaac	cttattcg	gtctttggga	35580
atgtaaagt	atgcaggcc	tttggaaaac	agtcggcag	cccttcaat	tattaaatac	35640
agagtaccg	tatgacc	gaatatttct	cctgggtcta	taacaaaaaa	aatgaaaaca	35700
tatattccaca	taaaaaactt	tacatggca	ttttagcata	cattattcat	aacagcaaag	35760
gtgttaagaa	cccataatgc	catcatctga	tgaacaggta	aataacatgc	gttattatcc	35820
atacactaga	atattatctg	cccatacaag	gagtgcacat	cagctatcat	ctacaaggat	35880
gaatctcgga	aaccttgc	taatgtggaa	aagccatgt	caaatacgca	cagattatga	35940
ttccatgcat	cggaaatgac	cagaatagg	aaatctatag	agacagaaaag	tagattatgt	36000
gttgggtggg	gctgggagga	caggtagtac	actactttt	cagaactact	ggaacaaaag	36060
accacaaaact	ggggagctt	aacatagaaa	ttgatttctt	cacaggctt	gagacttagga	36120
ctctgagatc	aagggtgtc	cagagcttgg	tctttcttgc	ggcccttgg	caaggctctg	36180
tcccaggct	ctcttcttgc	ctggcagggt	gccccatctt	ccctcg	tccatcatc	36240
ttttctctgt	gtgtgccc	gtccaaattt	tgatggctt	attctgg	atggccaatt	36300
gctatgcaca	aagtgaatc	tacttccaaa	agaaaggaaag	agggaaacact	gactaggcta	36360
aacttatagt	cattttatgc	tccgctttt	ctatgagatt	gtgaacacac	agaagtaggg	36420
tttttatcta	cattgtc	aaagttataaa	gaaaatata	attcaagaga	agcagttcaa	36480
tagcaggaat	ttaatatgg	aacttattac	aaaggttttag	gcaggactaa	aaaggccatgt	36540
gggatggta	gccaacccag	agatttgc	cagtgggacc	ccatcttac	accacccatg	36600
aagcttggaa	gataaaagg	gggttattat	cagagtccac	aagccatgt	cagagtcc	36660
ggctggagct	gggaccaccc	tagagacact	gtgcaaaagca	gaaaacaagg	gggaaaaacc	36720
ctgacttctc	ccttcttccc	accttcaat	ctcccaact	tgcttcttac	tagccatact	36780
tggccagaga	cagtgcac	gaacacttgc	aatatgtt	tgttagaaat	atctccctt	36840
gagacagaga	aatatgg	ggtgaaaat	ccagggat	atggatc	atggatc	36900
gtactatctt	atttatttctt	gtatcttgc	gaaatgtt	gtatcttca	aaaggaaag	36960
aatttggaga	aactgaaaac	tccaaattgaa	atgaaagaat	ggagaatttac	tggacttagaa	37020
gagaagagaa	aaatttattt	cgcataatgt	aaacaagaat	ggatttccaa	aggacgtgt	37080
gaatgaaaag	ctataatctt	caaagattt	ccagggat	taaaaaatgtt	ttaactcagc	37140
cacgtgtac	aaccttgc	cacaatgt	gaaaatgtt	caagaaatgt	caagatgtt	37200
agtcaaattt	taatgtctt	tccagaatct	ctcaagacga	ttatatact	acccat	37260
attaaataaa	atggaaactt	actaaactt	ccccttgc	taaaactaaca	tatgtcctaa	37320
tagcaaacga	ttcttgcattt	ccttagat	aatatattt	gtcaaaatgtt	attgtctt	37380
taatattctt	ctgaccctt	tttgctt	aggatattt	tatacatac	acacgtaaat	37440
ttggctata	gttttacatct	acgggtt	actgttctt	ttttcatttt	tttttttttt	37500

72/122

ccaaacccca	gtatccatat	actgctctct	atcagggtta	ttttaacttt	gtaaaatcg	37560
ctgagatgct	ttccatgttt	tttttttta	tttctgcga	catttgaata	gcataggagt	37620
taccaccatc	aaccttggat	tatthaagca	ttcacgattc	cacgtgtgga	tttttattc	37680
agatcttttc	ttgtcattcc	tgctatcagc	acagaaccca	atctcagctt	tccagctata	37740
ctctcaccatc	atggaaatttg	cagatgaagt	tccaaaggac	cattgcatta	tcctgcctcg	37800
ccctcttccc	ccttcattta	gacatcacct	tcttctagaa	cgtttaacct	gacatgccct	37860
gctcccaacc	cctgtgtgcc	aattgtgtgc	tctccctgt	cctggccctgc	cattctcttt	37920
agtaattgcc	tgctccctca	tctgtctccc	cacccagaca	ttaagctgaa	tagactggat	37980
ttgtgtcttg	tccatcaacta	ataatcgc	acttagtacc	tagtaggtac	ttaccatgt	38040
ttcatttagca	aatatgttatg	tataaccttg	caccttaaaa	acaagagaag	gaagacaaaa	38100
ttaagtctta	agactatggt	ttagaacatg	gatcagaaac	tacagtctgc	agcccaaatc	38160
cagaccaaat	gaagagagcca	tgttcattta	catacaacct	atagcgtt	tcacactaca	38220
ggagcagagc	taagttagttc	caagggaaaa	cacggccctg	caaagctaa	aatatttact	38280
ctatactct	tacacaaaaaa	agtttgcaga	tcctcggtt	agaactcttg	ttcatatgt	38340
atttcactacc	accatagttt	tttgggttg	tttggttttt	tttggcaaaa	aggaatgagc	38400
cgatccagaa	aagggtgaaa	agaatgaatc	attactgctg	aaagaatgtg	cacacagtcc	38460
gtcagtattc	tgctgccccat	ctgacaccca	tccaaatagtg	tcatggatg	cagcagctac	38520
tactgtgtt	tcaatggcca	gtccacccac	tcctataacca	tgtccaagca	atcttggaa	38580
catcatcacc	atgttgtt	atcctttaagg	tattgcctca	catacagcag	tggctggta	38640
taaagtcaaa	tgacactagt	ggccaggagg	tcaagagaat	gagttagggac	aggtgggtag	38700
gcagccccagg	ccctagcaac	agcaggagct	cacccctcg	tcactctagc	caggactgaa	38760
atacttttca	ccctttcaag	agagactagg	aatctggatt	tttatgtgaa	atatcttgc	38820
tactaaatgt	tgtcaacaga	catgtcaaaa	gttaaaacta	agtaaggttca	tggggcagat	38880
tgacttattca	ggtttagaa	ttaaggattc	ttagccaaaca	cagataccaa	ccaaaaagct	38940
gacgtataac	atattaggag	aaactatgtt	cactgtcgaa	acatcaacaa	ggggctaattg	39000
tctaaaatag	tctatatttg	attccagttt	aaacatgggg	aaaggacatg	aacaggcaac	39060
ttatgtcaat	ggaaaactcaa	aaagataaca	agcataatata	aaagcatttct	caaatttgc	39120
agtaacacaa	cagatgcaaa	taaaaagagg	aaaactgtctg	ccgggcacag	tggctcacac	39180
ctgtatcccc	agcactttgg	gaggccgagg	cgggcggtt	atgaagtctg	gagatcgaga	39240
ccatctggc	tcacatggt	aaacccctgc	tctactgaaa	acacaaaaaa	ttagccaggc	39300
gtatgggtgg	gcaccagtag	tcccagctac	tcaggagtt	gaggcaggag	aatggcatga	39360
acccaggagg	cgaggatttc	agtggccga	gaccatgcca	ctgcactcca	gcctggcga	39420
ctgagtgaaa	ctccatctca	aaaaatataa	tataattatt	aatattataa	ataataaaata	39480
gtaaaataaaat	aaaaagagag	agactgtca	agtcttagaa	gttgaatgt	gccaagcgc	39540
tgcaaagatc	agggcccttg	gatggccggg	tgcaagtggct	cacccctgt	atcccaccac	39600
tttgggaggc	caaggccggc	ggatcatgag	gtcaagagat	caagaccatc	ctggccgaca	39660
cagtggaaacc	cggtctctac	taaaagtaca	aaaaatata	tatataatata	tatatttata	39720
tattatataat	atataatata	gagccctggg	aatcttgcgt	tgctgtggg	gaaggttagt	39780
gtgcagccac	ccttgacagc	aatcttgcgt	tacttggta	tatataat	aggcacacac	39840
cacgaccagg	cagtcttact	cctgggtct	aatcccaaag	aattcttca	caagtccata	39900
aggagacatg	tacgaggctc	atttcagatt	actgggagtg	ggaatcaacc	tgggtgttca	39960
tctacaggag	acgagatgga	caaaatgtgg	tggatattaa	gaccagaatc	accaagtaac	40020
agagatgggt	ggtgagtgt	aatccctaaga	tacagaataaa	aggctagaac	atgtgcct	40080
tcatgttaat	taaaaatata	tgcacacaaa	cgatatacg	cgtgaccctt	gaatagcaca	40140
ggtttgaact	gcctgtgtcc	acttacatgt	ggatttctt	ccacttctgc	tacccccaag	40200
acagcaagac	caacccctct	tcttcctct	ccccctcagg	ctactcaaca	tgaagatgac	40260
aaggatgaag	acttttatga	taatccaatt	ccaaggaact	aatgaaaatg	atattttctc	40320
ttccttatga	ttttctttat	ctctagctt	cattattctt	agaatatgtt	acataataca	40380
catcacacgc	aaaataatgt	ttaatttact	gttttatata	tgggttaggc	ttccactcaa	40440
cagtaggtgt	tcatgttttt	agttttggg	gtaaaaggat	atacacatgt	tttcaactgt	40500
gcaggcaatc	agtttccctg	accccctcat	tgttcacggg	tcaactgtat	atacacaaaa	40560
gtattatatg	aacctcatta	gaatagctgt	ctataggggag	aagagaatga	gagtgggata	40620
aaacggaaatg	aacaaataaa	ccaaacaaatg	cattaacaaag	caaacacaa	gagggggcttg	40680
catggggccag	tgatgataaa	gggctaaagaa	tgagaataaa	attaatcaa	ttcctcacac	40740
ctgagggtcta	aaaccaagga	aaggggggc	caggcgtgg	ggctcacccc	tgtatcccc	40800
gcactttggg	agggctggagc	ggggccatca	caagatttgg	agtttggat	cagcctggcc	40860
aacacagtga	aaggccatct	ctacaaaaaa	tacaagaatt	acccagggtt	gttggcacat	40920
gcctgttagtt	agctactctg	gaggctgggg	caggagaatc	acttgaaccc	aggaggccga	40980
ggttgcaggg	agccgagatc	acaccattgc	actccagcc	gggtgacaga	gtaaagact	41040
gtctcaaaaa	aataaaaaaaa	ataaaaaaac	agagaaagg	aggaaactag	atccaggctg	41100

73/122

actagataca	gccttagag	ttagaaaaga	tgatttaca	atctaagccc	acactcgat	41160
tgaatgaaat	tgaaaagcct	ttcaactaa	aacatttaat	tacaccatct	gtgcagaca	41220
gaactcagac	aactcaaaca	ggtaatgtca	gcgtgggtt	ttatatcacc	accctcaaca	41280
cagaataaaaa	atcagctgca	tgtgaagcag	tgactagaat	gaagaaaagg	ctgcttccta	41340
cttccttcta	gtggttctt	ccgaaaacat	taataggcac	cagctctatg	catgtcaccc	41400
tgcagggaga	catggggtat	ataactatga	tttactgttc	attcctcaag	gaattccccaa	41460
tcttgtggaa	gattatacac	aatgaggca	caaaaactat	ccaataaaaac	cacggaaaaag	41520
aagccagtga	caaaaaggc	agtgtgaa	ggccctgtga	gcagagctga	tgccatttg	41580
gggaagaaag	accaacatgg	atgggggtga	tcaggggtgc	tccgtggaa	agcttggaa	41640
gaagtggcag	atctctgagc	tggatgtatgg	gccactacca	tctgtatatg	gctaattaaa	41700
gaccatgtgt	ggattttta	ttcagctctt	tcgtgtcatt	cctgcttatca	gcacagaacc	41760
caatctcaac	tttccagctt	tatttagatca	aacttctcac	ctcatggat	ttgcagataa	41820
agttcaaaag	gatccttgc	ttttcaaaaat	aattttgaat	ggtttagat	tccctctgtg	41880
ctcttcact	gacaccctct	caaggctgct	gagcacgtgc	catgtctatgg	ctttctccaa	41940
catcaggaaa	tgttctccac	tcagttcac	cttaataacaa	atgtttctc	tcttcagaga	42000
aggcaaaaaa	attcatgacc	atctgactgg	gagaagtcat	ttcttaggtaa	agtgtccatc	42060
ttttcttgcag	gaacacagga	ggaaaatctt	acagaaaaga	gttaacacag	caggcctaag	42120
actgctttt	aaaataaaaat	aaaataaaaat	aaaataaaaat	aaaataaaaat	aaaataaaaat	42180
aataaatgaa	tgatagggtc	ttctgttattg	ggccaggctag	tctcaatttc	ctggcttcaa	42240
gagatcctcc	caccttggc	ccccacagtg	ttgggattat	agacatgagc	cattgtgtctt	42300
ggcccaagac	tgttattctt	aaaaagtc	ataaaaagca	tggtaatcc	ttggctggca	42360
cctggaaatc	tagatttcag	aagggttccc	accatccaaac	ctggaaagag	ggactcaactg	42420
tgccctaaatt	atttgtgtt	ttatgtctgaa	ctctctgttt	tcttcaggta	gcgttggaaatg	42480
tggatgtgc	tggccaaagg	ggccctgc	gaccagcccc	caaaaaaaaa	ccttgggttt	42540
gggttcttag	tgagtttccc	ttgttagacag	catttcacat	gcgttgcac	agcttcttcc	42600
tcggggagtt	aagcacatac	atcctgtgt	actgcactgg	gagaggatgc	ttgaaagctt	42660
gtgccttgc	tcctttggac	ttggccccat	gcaccttcc	cttgcgtgt	tgtgttttgt	42720
atcccttcac	tgtaaaaat	taacagccgt	agtacaccac	atgcgtgt	tcccaagtga	42780
accaccatg	ctgacatgg	tcctttgggc	cccccaacaca	gaaataattt	ataaaagacc	42840
aaggactggg	catggggcc	catggccgt	atctcagcgc	tttgggaggc	cgaggcagga	42900
ggaccagtt	agccaaaag	ttcaaaagtt	cagtgcacca	tgactgcgc	aatgcactct	42960
aacctgggag	acagagcaag	accctgtccc	caaaacaata	aactaaacac	atacttctgc	43020
cttccaagtg	tctaaaaatt	caatggat	gtagaaacat	ttttaaaaca	ctaaatcaaa	43080
agaaaacctgg	aaaacaagag	tgcgcgttgc	caactaaaaat	gtcttagaaaa	tttctgaaaa	43140
gtaaaaagta	ctcagaacca	gattacctg	gcaaaaccata	gccccatata	agcttgggag	43200
gaggctgtt	tgcagaagga	aatggtaaca	gtttccagg	aacagacttg	taacagcaga	43260
tagaacagca	gaggttaga	ctgacaaagg	gattacctg	ggaactgcag	tctgaatgac	43320
caggactgtt	ggacccttcc	cctcacaatgg	aatacacac	ccactcagca	gcacaccaca	43380
gtcttcac	aatcaca	ggcgcgttac	gccttagtaa	acaggaaaaaa	agaattctc	43440
aaacttcgaa	gtgaaacaca	taaagaatca	ccaaatttt	attcagat	atgaaacagg	43500
gacactgaat	caacagaaca	caaaccacaa	caaagataat	tactagagca	catagaagaa	43560
attattat	attcttgg	agacctaagg	ggacattata	aagagca	atgttggat	43620
tgacatctt	tgtgtat	caagaaataa	aaacacagg	tgaagacc	atagagaata	43680
atgtactat	tttgtcaaa	aaggaaaaat	ggagaatctg	attcatattt	gtttgtat	43740
gcataagaa	atcttggaa	gtacataat	aactaacaac	aatggtacc	tacttgcata	43800
gcgagagaag	taagaggaca	ggaatggtgg	gaacaccctt	tgtgtccgg	attgggtgg	43860
tcttggctcg	acttggagaa	tgaagccgt	gaccctcg	gtgagcgtaa	cagttctt	43920
aggccgtgt	tctggat	gttccctgt	atgttggat	gtgttccgg	tttcttcc	43980
ctgggggtt	cgtgtctcg	ctgactcagg	agtgtacgt	cagacccctcg	cgccgaggt	44040
tacagctt	aaaggggcgc	atctagat	gttgcgttcc	cctgggt	tctgggtctc	44100
gctagctca	ggagtgaagc	tgcagaccc	cgagggtgt	gttgcagctc	atataagacag	44160
tgcagaccca	aagagtgc	agtaataaga	acgcattcca	aacataaaa	ggacaaaacct	44220
tcagcagcgc	ggaatgcac	cgccacacgt	taccactt	ggctcgggca	gcctgtttt	44280
atttcattt	ctggccacac	ccatatactg	ctgattgg	cattttacag	agagccgact	44340
gctccattt	acagagaacc	gattggtcc	tttttcagag	agctgatgg	tccattttga	44400
cagagtctg	attgggtcgt	ttacaatccc	tgagctagac	acagggtgt	gacttgggt	44460
tttacaatcc	cttagctaga	cataaaagg	ctcaagtccc	caccagactc	aggagcccg	44520
ctggcttcac	ccagtggatc	cgccatcagt	gccacagggt	gagctgcctg	ccagtcggc	44580
gcccctgc	cgcaactcc	agccctctgg	tggtcgtat	gacttgggc	cgccggc	44640
gggggtgggtc	tgtcaggag	gtcggggcc	cacaggagcc	caggaggtgg	gggtggctca	44700

ggcatggcgg	gcccgcaggc	atgagcgctg	ccccgcaggg	aggcagctaa	ggcccagcga	44760
gaardatgggc	acagcagctg	ctggcccgagg	tgctaagccc	ctcaactgcct	ggggccgttgc	44820
ggggccggctg	gcccggccgt	cccagtgcgg	ggcccccggca	gcccacgccc	accgggaact	44880
cacgctggcc	cgcaaggcacc	gctgtagcagcc	ccgggtttccc	cccgccgctc	tccctccaca	44940
cctccctgc	aagctgagggg	agctggctcc	agccttggcc	agccccagaaa	ggggctccca	45000
cagtgcagcg	gtgggctgaa	gggcttcctca	agcgccggcca	gagtgggcac	taaggcttag	45060
gaggcaccga	gagcgagcga	ggactgccag	cacgctgtca	cctctoactt	tcatttatgc	45120
cttttaataa	cagtctgggt	ttgaacactt	attatcttac	ctattttttt	tttttttttt	45180
tgagatggag	tegctctctg	tegccccagac	ttggatgtca	tggtgcctat	ctggctca	45240
gcaagctccg	ctcccccgggt	tcacaccatt	ctcctgcctc	aacctcttgc	gtagctggga	45300
ctacaggcaa	tcggccaccac	gcccaagctaa	tttttttattt	tatttttttt	tttagtagaa	45360
cggagttca	ccatgttagc	cagatggct	caatctctgt	acctctgtat	ccatccgcct	45420
cggcctccca	aagtgtctgg	attacagacg	ttagccacttgc	cggccctgcct	atcttaccta	45480
tttccaaatgt	taaactttaa	gaagtagaaaa	ccccgtggcca	ggcgtgttgc	ctcacgcctg	45540
taaccccggc	taacccgggg	gcccggccgg	gcccgttgc	aggttcaggag	atcgagatca	45600
tcctggtaa	cacagtggaa	ccccgtcgct	actaaaaata	caaaaaattt	gccggggctg	45660
gtggggggca	ccggcagttcc	tcgtacttgg	ggaggcttgc	gcaggagaat	ggcgtgaacc	45720
tggaggcag	agcttgcagt	gagccgatag	agtgcatttgc	ccttccagcc	tggcgcacag	45780
agcgagactc	caccccaaaa	aaaaaaaataa	aaaatagaga	ccccggaaatg	taaaaatatg	45840
ataatataa	ttttaaaaac	ctcaagagat	gggtcaaaaa	gttgcaggaa	caaataatca	45900
tatttagtt	gtgcacactg	aaaccaggccc	aaggAACATC	ccagaatgc	gccccataaaag	45960
ataaaagagag	cattttccgt	gggcacagt	gtatggcagg	ggaatttgcct	gagtccaa	46020
gttgcaggtc	acattgaacc	acaccattgc	actccaggcc	tgggcacac	agcaataactc	46080
tgtctcaaaa	aaaaaaaaaa	ttaaattaaa	aaagacagacaa	tatttgcagg	aaaaaaaaatgc	46140
ttatttcaag	aaacatggaa	gataaatcaa	gatattctaa	tttcccaatg	agaataattc	46200
cagaagcaga	aaataataga	gagggcaagga	aacactcaaa	acttctccag	tgccatagaa	46260
atgtgtatta	atcttttagaa	tgaacggac	taccaatgc	tgagcaggaa	gaacaaaaga	46320
gatccactct	taagccagt	ttgtgtccaa	gcccgttgc	tcatgcctgt	aatcccagca	46380
cttggggagg	ccgaggcagg	ttggatcact	gaggttcagg	gttgcaggatc	agtccaggcca	46440
acatggtcaa	accctgtctg	tactaaaaat	aaaaacattt	gttgcggat	gttgcacaca	46500
tctgtatcc	caactactt	ggaggcttaag	gcaggagaat	cacttggaaac	caggagggtgg	46560
aggtttagt	gagccgagat	catgccacac	tcccagctgt	gttgacagag	caagattcca	46620
tctcaaaaaa	aaaatccact	ccttagacaaa	taatagttaa	attttagaa	accaaggaga	46680
aagaaaaaaa	attgttaaagc	ttcaagaaaa	ataaacattt	actacaaaa	aacggagatc	46740
agacgcgtgc	acttcttct	agataccaggc	agataaaagca	atatctccaa	aattcagaag	46800
gttttaacgt	agaatcttat	accaggctaa	gaatattc	atggaaaaatg	gaataaaaaaa	46860
acattgtta	aacatgc	ggttcagaaa	gttaccatt	cacagaatcc	ctgaaaacaa	46920
aaccaaaataa	tcacttaagg	actcattaag	aaaacaaatg	aaataaaagc	accaatgtat	46980
agtaaataat	cagaaaaatt	tgatgttac	ctaaataact	gttgcattgc	aatgtatgaa	47040
aacccaaaaa	tttataatgg	gacagaat	aaatcatgt	aaatgtat	tttgcatttac	47100
tcatggagag	ttcacataaa	cagattatct	ttaatagca	agagaaaaaa	atgttttagat	47160
atgtgtaaa	aactaaagggt	acccaaacag	tgcaatttca	tttatcatca	gaaaaatcca	47220
aattaaaacc	acagtatcca	ccagaataac	taaaaggtaa	aagacagaaa	ttaccaagag	47280
ttggcaagaa	tgtggagca	ccacatatac	ttctggggta	aataagggttgc	tgcacccgggt	47340
actgaaaact	gtttgcgtat	atctactaa	accggacaca	tgcacagact	acaaccaagc	47400
agttccactc	ccagatacac	actcaacaga	aatgcacaca	ctcacttac	aaaagacgtg	47460
tactagatgt	ttcatgtact	tactattcat	aatgtccaa	aatgcacac	aaccaactgc	47520
caatcaaagt	caaataatata	tctatattag	ggatata	aatggcatat	acacagcaat	47580
gagaatgaaa	tgaaccagtc	cgccacatgt	tttgcattgc	gtatctcag	cacttgggc	47640
gggtttagggca	ggcagatcac	ttgaggct	aaatttggaa	ctagccgttgc	caacacgggt	47700
aaaacccatgtc	cccaactaaaa	acacaaaaat	tagccggca	taggttgtgc	aggcctgtaa	47760
ttccagctac	tcggggaggct	gggttggggag	aatcgtttga	acccggaaac	cgagggtcgc	47820
agttagcgg	gatcgtgca	ctgcacttca	gcctggacga	tagagcaaga	ctccgtctca	47880
aaaaaggaaa	tcaaaaatat	aaaataatgt	gacaggaaata	atccgcacaaa	gtcagtaat	47940
caaaataaaat	ataaaatgggc	taaagctacc	tataaaaaaa	caaagatttc	acacccatata	48000
ggatagctac	tatcaaaaaaa	agagagagaa	taacagatgt	tagcaaggat	gtatggaaac	48060
tgaatttctc	acgcatttgc	ggtgatgaaata	taaaatgggtt	cagcctctgc	gaaaaacact	48120
atgttgggtc	atcaaaaaat	taaaataga	agtactactt	gatccaaacaa	tttacttct	48180
gggttatatac	ccaaataact	gaaagcaggg	tcttgaagag	atatttgcac	accatgatc	48240
atggcagcat	tattcataat	agctatgtat	ttggaaaccaac	ataaaatatcc	tttgataat	48300

75/122

atatggataa	gcaaaatgtg	gtgtatacat	tcaatggaat	attaattagc	aataaaaatg	48360
aagaaaaattc	tgacacatgc	tacaacatgg	atgaaccttg	agggcattac	attnaatgaa	48420
ataagccagt	tataaaaaga	caaatactat	atgaggact	atattagata	ctcatgcaag	48480
gtacctaaaa	taggcaatt	catagagaca	aaaagcgaaa	tgggggtgc	caggggctgc	48540
ggtatggat	acagagcttc	aattttgtaa	gtgaaaaaaa	ttctggagat	tgggtgcata	48600
acaatgtca	cacacttaac	actggggAAC	tgtaaacttA	aaagtagtaa	atggtaaaaaa	48660
taaaaataat	aaataataaa	tttatgtta	ttttaccaca	atatttatta	aaagacaaaag	48720
attaactaat	taaacaaaat	ccagccataa	gctaattggta	agagtaacaa	ttaaagaaga	48780
cacagaaaaat	tgaaaatcag	tgacttagaa	agatatttc	atataatgc	taacaaaaag	48840
caagtacagc	aataaaaaaa	gaatgacaa	aaaaaaaaat	aaataaagtc	gctcggttat	48900
tcccgggg	tacaatttac	caagaagata	caagaattgt	gaacctttaa	gcacataaaa	48960
cagcttcaaa	aataacaacat	ttaaagaaaa	atatatatta	aacatagaaa	tagtacaaaa	49020
acccttacaa	gaatcataat	gggagtcttc	aatacaactc	tccatataca	caggtaaacc	49080
agagaaaaaa	aataagttaa	ggatgcagaa	aacctgtaatt	accatcaata	aacttgagat	49140
taatataaaaa	ctgtatacc	aataataaa	gagttcagg	aacatgcgt	actgacagt	49200
gactgcaat	taatgttc	ttaatcttgc	tttttcttcc	agcactgtgg	cagaatagag	49260
atctttaaaa	ccttccagct	acaaaacatc	tttttaaaaa	tataaaaaaa	tacaaaaata	49320
actctgaaat	caatagaaga	cacatggta	aaccaaaatt	ctagaataca	gggagaataa	49380
aggcattttc	agatatttaca	aaaacagaaa	attgtatcatt	gctgaagtaa	tttctaaaga	49440
atgtacttgc	ggggagaagaa	aaatgttcca	agaaaaagta	tctgtgtatc	aagaaggaaat	49500
ggaaagtggaa	gaaatggtaa	acaggtagat	aaagctaata	aatgttgc	tagaaaataa	49560
caaaaacaat	agcaataatg	tctcggttga	aggggttgaag	taaaatatac	attaaggcca	49620
aatgtgaggt	aagtggaaatg	aaagaattag	aagtccctgc	cttgcataca	ggactgatta	49680
aataaatgag	ccagggtttc	cattcaaac	gttaaaaactt	gaacaaaatc	aactcaaaatt	49740
aatgatggaa	ataaaaaaaca	gaaattatgt	tcatagaaaa	ataaaaatc	aatagaatta	49800
atcaataat	cctggtaat	aaaagctgtt	tctttgaaag	gattaataaa	ataatcatta	49860
agcaagtctg	atcaaaaaaa	aagagaaaaag	gtaccaaaaa	aagtactgt	tcagaaagag	49920
aacatacaga	tacatacaga	tatgtttag	tctgtttct	tacaccagaa	tactatatac	49980
aacattatgc	tagcatat	taaatttcaa	taatgttaat	gattttctag	gaaaacagaaa	50040
aatattaaat	ttacttttga	gaaacagaaa	aactggaaaa	aataatgtat	catgaaaaaa	50100
atggaaaagg	aattaaataat	tgatattaac	tgccttaaac	acaccagcg	cagcccaggc	50160
agtcgtcagt	caagttctgc	caaacttgc	ggaacagata	attcttctat	tccagagcat	50220
agaaaatgt	ggaaagtttc	ccaaatttaat	cagagaggac	agcctgatcc	ttgttatgaa	50280
cacagataaa	atggggtaa	actatatgcc	aaactcgat	accaaaaacc	taataaagat	50340
gctagtttat	tgatgtgaac	aatccaaag	tgcattttaa	attagcccg	ggttttagag	50400
aaagaaaatc	tagcaatgt	accaccactt	atgttaacaa	ttttaagacg	aaaatctaca	50460
tgatcatatc	aatgcatact	acacaaaagc	atttgggcaaa	aaaaccaac	accaccctt	50520
gactttttaa	actcttagta	attaggcata	aacagaaatg	tacttaatgt	gatagaatac	50580
actcggtgaa	gatacagagg	gaatgtccc	tttttttttt	ccccacaaa	agattccat	50640
ttaacccatca	tagtcaacac	tgcagcgaga	gtatctatg	gaagacaagg	aaaaaagtaa	50700
aaacatgaga	gacatctgtt	gtttaacaga	caataagatc	acctacttgg	aaaggaggaaa	50760
cgaatcaagc	gaaaaactat	ttaaacttgc	acaggctttt	gtatggaggc	tcagttcag	50820
ctgtatgttt	ggcttacaaa	ttcaacttgc	ttgttggag	agttatctt	gcaaagctaa	50880
tttctgttga	ggtatttaga	ttgacaagcc	tgtgtccccc	cctccccc	cattttcaac	50940
actggaaataa	cacgggtttt	ggaactggat	aacagaatct	tccaaaaaca	aaaattgtcc	51000
tgaaggggctg	acttggcc	ttactcaaaa	aacactttat	ctgctgcctg	cagctcctac	51060
agttgtctgt	ggataaggct	gcccaaccgc	tccgcgttaat	tcttctgtca	gagggcaagg	51120
aagagactt	tcacagggaa	attttttcc	gaactgtatg	ccgcttatta	cataaaactta	51180
cgtgtctggca	aatggagtc	cagaaaaata	agatatttcag	agtccaaactt	cctttaggaaa	51240
aaaaaaaaaa	aaaagcaagc	acataacact	attttcttg	catggcact	ggggaaaggag	51300
gtcggtactt	ccgcacccc	gcagggtccgc	accaccggga	aaccacgggg	caccgcgcgc	51360
tggccccccgg	ctttccaggt	gcactgcgc	ggggcgcccc	agctgaccgg	ggatgcgcag	51420
ccctagccct	tccccctgtca	ccccggccag	gaagggggcg	gagcgcggcg	gacgcccgg	51480
gcaaggggct	tctcggtccct	ctgcacccac	cagcacccccc	aaggcacaac	agggagggtg	51540
cgggaggctc	ccgagaccca	ggagccgggg	ccggggctgc	ccgcgcaccc	gtcccactgc	51600
ggcgagggtct	gggggtcgcc	ccaggccgc	agctgtcg	agccacccctg	cttctcgtcc	51660
cggggtccctg	cgacaaccc	cgggccccga	gggggaggagg	cgccacccctg	ccgcgtccac	51720
ctgcggcacc	ggtcccaccc	ctccggcccg	ggcaggacag	gccaggacgt	ccctcttggg	51780
ctggggacac	gacacgcgac	gaggggaccc	ggggccccccgc	ggcgaagacg	cagcacccct	51840
tcccagaaag	gcagtccctgt	gcccccaacga	cggaactgcgc	gaccccccgc	ctcgccccgc	51900

76/122

catcccttca	gaccacgcgg	ctgaggcgca	aagagccggc	cggccgggcgg	gctggcgccg	51960
cggttagtac	tcacccggcc	cgctggctca	gcgcggccgc	aaccccccagc	ggccacggct	52020
ccgggcgctc	actgatgctc	aggagaggga	cccgcgctcc	gccggcgct	ccagccatcg	52080
ccgcccgggg	gwgagcggca	gcccgcgggg	gtcgctgg	agatgtaga	cccggaaccgc	52140
cgcctgcgcc	gtcctccccc	agccggggc	ccccggccccc	ctctctccca	gctctcagt	52200
tctcatctcc	ctatctgctc	atcccttgtt	cgcacataat	cgatgtttgg	gcgtcccaag	52260
ccagatgtgg	accccatttc	cgcactctac	actggaggtt	ttctaaagggt	ggtgcccgg	52320
ccagcagctt	cagcctcatac	tgggaatttg	agaaaaatgc	gatttccgt	cccacccagc	52380
ctattegggtt	tttctgtca	aaaaccatg	aaagggtgggc	ccagcagtcc	acattctcg	52440
aagcccgctca	agtgatttg	aggcccttc	cagtttggaa	gctatgtca	ccgcctcacc	52500
tccgccccgc	aaggagcccg	gtcttgcctg	tggcgctagc	cgcacacgg	cacctcatcc	52560
tgcggggccc	gccccccccc	tgaccaccca	ccggcccaacg	cctcccccgg	gatgcagcgg	52620
aggcgccctgg	aagtccggca	ggtaacatc	ccccctcagca	tcttccctac	cctcacggct	52680
cctccctecag	gggtgcetca	tggccagggg	tttagaaagag	ccactgtgtt	tcttgacatg	52740
gaagtggccct	aagacttta	tgaaaactgc	aggagtgaaa	tgacagaacc	tttggtcata	52800
cttgaggggcg	tgaagctcaa	atgaggagga	aggaaaaggat	ccaggagagaa	taaccaaccc	52860
tggcaagttt	tggcgccca	gtagaggggc	gagcctaggc	tagcggttct	cgaccagggc	52920
cggtgttgcc	cctccctcgcc	gcccgcgt	cattttggga	ggtctggaga	cattttgggt	52980
tgtcatgat	cgggagttgc	tacttttgcc	taatgggtt	gacacgaggg	tgctcctcaa	53040
catcttaccc	gaaggacagg	actgccccac	aaggaagaat	gatccggccc	caaataagaa	53100
accctgggct	ggtcagcaac	aaccctttt	ttctgagaag	agaggagggaa	agaataaaaag	53160
aagtgggtt	aagtttttgt	ttgttagagg	aaacttgaag	acattttcac	tggaaaggaa	53220
gagaggaaga	ggagggagat	gtctgtaa	acgagcaac	cggggtacag	ctgatttcc	53280
catattgaag	taatgatcc	tagttataat	aaatttctaa	ttaaaaccca	ttttatccct	53340
gcaataaaact	tgtttttttt	ttttaaatat	actgtttgtat	tctgtttgt	aatattttat	53400
ttcaggcgtt	tgcatgtata	tgcaaaaatg	agatggca	taattttctt	tttgaatgtc	53460
taatgttgtt	tggtttca	atcaatgtt	tgctcacatc	ataaaaaaatt	tggaaaccgg	53520
gcaggaggag	tgcttgaggc	cagaatgtcg	agacaggtct	agggaaacaca	gtgagacccc	53580
cccatctcta	aaaaaaaaaa	aaaagaaaaaa	aaatgggca	tgtttcttt	tttcttttac	53640
tctgaacaat	ttaaggagca	ttaaaattat	ctattttttt	aggtttgtac	atttcccagt	53700
taaaaatgtt	cctcccaagcc	tgatgtttc	tttggggagg	gtaaatctt	taaggctaga	53760
aaagttctt	ctgtggcaat	tttatttattt	acattttaaa	aattattcta	gagttattt	53820
tgataaagca	tgttattttt	aaaacaattt	atcctttttt	tccagatgtt	caagtgtatt	53880
tgcataaagg	tgagggaaatg	agtctttttt	gaatctttt	acttctccca	aatatcttat	53940
tttgttattt	tttgcctttt	tattttgtt	actttttaaa	gtgttattttt	ttttcaaaaga	54000
atcagcttt	aggtttatgt	ttttgggtat	actggagctt	ttttcttctt	ctttttaaa	54060
tatttttctt	cctttttttt	ttagacgtat	tttgatctt	cgtaatcg	agaaggtaaa	54120
tttagaatctt	ttgttactat	tgtttttttt	tttctcttca	tttctctgaa	gtcctgttt	54180
ataaaatgtt	ccatgttat	tgtcataaaa	tattttttt	tcttattttt	tttgcatttt	54240
tccacttca	tcataaaaatg	accttcttg	tctcattttt	tgtttttttt	ttttcaaaaga	54300
aatttaactt	tgtctgatat	ttaaccatcc	tgctgaattt	tgtttttttt	cccaaacaac	54360
cttgctgtt	ttcgtttttt	ctgaaccctt	tattttaggt	aatcccttga	attagagcac	54420
taagtttgc	tttgcattttt	aatttggaaa	tctttatctt	gcccataatgt	atgttgcggcc	54480
tattcatgtt	acatgttat	tatgtttttt	catagccctt	tttgcctttt	tttctacttt	54540
gcattgcata	ttttgtttt	attttgtttt	gttgccttc	tgataatttt	gaagggttgt	54600
attttttttc	agggagggtgc	tttataatca	tactccgc	tacacatcg	cctcagttt	54660
ttcagactgt	ctgttaactc	cctattctga	ataaaaatga	cattttaatt	tccctttttt	54720
ttcttttaccc	cttttcttct	cctcacctt	tgttaatgtat	tttacccctt	tttagtattt	54780
gtttttttaa	ttaactacat	tttataatat	ctttatctt	tgtatttttt	atagctttt	54840
aatgagat	tttgattctt	agataaaaaa	gtgtttaat	ataccatttc	cacgtttagt	54900
gtttttttaa	atcatacatt	ctgtgtgt	accataatcc	cacgtttgtt	tttagttccac	54960
tcttacagtt	aaaagattca	gaagtattat	taacagtat	tttgcattag	ttttttccccc	55020
aacccatttt	gtggtaagtt	atgatccctgc	ttttagttctt	taagaataat	ttatagagca	55080
gagtgtggtg	gctcacgtt	gtatcccg	cactttggga	gacaagaggt	agaaggatcg	55140
cttgaaggcca	gcagttcaag	accaccctga	gcaacatgt	gagacccctgt	ctctacaaa	55200
aattttttaa	tttagccaga	cgtagtggcg	tgtgcctata	gtcccaagct	ctcaggagggc	55260
tgaggcaaga	ggattgttag	agcccagaag	tttgaggctg	cagtgacctc	tgattgtgcc	55320
actgcacccc	agtctggca	agaaagttag	aacctatctc	ttttaaaataa	caataataac	55380
ttatgaaaat	tatattccct	gagtttttca	tgtttaaaaa	tatttgc	tttctatcc	55440
taaaagtttgc	agtataaattt	cttgggttat	actttatcta	tttgcataat	tataagtatt	55500

77/122

gtcttctaga	attgagtgtt	gctgtaatga	aaccagaagt	cagcctgggtt	tat	ttttccct	55560
cagaatgag	gtaattgcgg	gccggacacc	gtggctcatg	cctgtatcc	caacactttg	55620	
ggaggccgag	acagggtggat	cacgggtca	ggagattggag	accatcttgg	ctaacatgg	55680	
gaaaccccg	ctctactaaa	agtacaaaaaa	gttagctggg	catgggtgg	gacgcctgt	55740	
atcccagcta	ccccggaggg	tgaggcagga	gaatggcgtg	aacctgggg	gaggagctt	55800	
cagagagctg	agatcgcgcc	actgcactcc	agcctgggg	acagagtgg	actccgttcc	55860	
aaaaaaaaacaa	aaaaaaaaaca	aagaagtggaa	gtaattgcctt	tgatgccttca	agaatttatct	55920	
ctttgtctat	gaaatccaga	aatctactg	ttatacattt	tggattattt	attctggggcc	55980	
aatatttct	gggacacaat	agattgactt	tagatgtttt	at	tttttttttt	56040	
acagagtctc	actgcatttc	cagcttactg	caacctctgc	ctcacgggtt	caagcaattt	56100	
tcctgcctca	gcctcccaag	tagctggac	tacaggcgcg	tggcaccat	cctggctaat	56160	
ttttgtcttt	tttagtagaga	cagggtttca	ccatgttggc	caggctgttgc	ttgaacgcct	56220	
aacctaagt	gatccacccgt	cctcagccctc	ccaaagtgtt	gggattacag	gcttgagcca	56280	
ccatgcggcag	cctcaatttc	tcttttattt	tggttattttt	tctgaattttt	aaaacatttgc	56340	
ttctaatcag	ttatatttctat	gttcttctaa	gatgtgtaaa	gcacccattt	cccaagggtcag	56400	
cccccatctt	gctagtggac	teggctgggtt	cttcacaaga	gctctggttt	tctcctgttt	56460	
aatctcaagt	acctctgtca	gcctccaccc	gttttatgt	ttggagttttt	ttgggtttttt	56520	
tttttgcgtt	ttgacagagtt	cttactctgt	caccgggtt	ggagagcgt	ggcataatct	56580	
cagcttactg	caacccctgt	cttccagggtt	tgagcgattt	tcctgccttca	gcctactgttt	56640	
tagtgggtt	tacaggcgcg	tgccaccaca	cccggttata	ttttgttattt	tttagtagaga	56700	
tgggtttca	ccatgttggc	cagggtggc	ttgaacttctt	gacccagggt	aatccacccgt	56760	
cctcagccctc	ccaaagtgtt	gagattacag	gcgtgagcca	ccgcgcctgg	catgggtttgg	56820	
agtttaatc	tgttagttttt	ataaagatag	tgcttatgtt	tgtgtttttt	atatttcttgc	56880	
gtactcttgg	gtatgttgc	atataccttgc	aaatccctggg	gaagtttattt	tcaattttttt	56940	
ttcttcagac	tttttattttt	attttattttt	attttattttt	tatgtttttt	atggagttttt	57000	
gctgtgtcac	ttctggaggc	tggagtgcag	tggcgcgtttt	tcaggttact	gcaaccccttgc	57060	
tctccgggt	tcaagcaattt	ctcctgcctc	agcctcccg	gtagctggga	ttacaggcact	57120	
ctgccacttt	ttaattttttt	tagacacaga	gtctcgctttt	tttgaccagg	ttggagtttttt	57180	
gtggctgaaat	catggctgac	tataaacttgc	aaatccctggg	ctcaattttttt	tcctctggctt	57240	
cagccctctg	agtagctggg	actacaggca	catggccacc	tgcccaattttt	atttttttttt	57300	
ttttttagag	acagggtttt	catatgttgc	ccaggctggc	ctcctactttt	tggcctcaag	57360	
taatccctct	acctcagcc	cccaaattac	taggattata	agcatgagcc	accatgcccc	57420	
gccttgcctt	actactttaa	tttcatatgt	taggttgcac	tgtatgtt	catccaaacc	57480	
aggataatgt	aagaatgaaa	gagggtgaca	gttagatgt	gttggactt	gtatttttttt	57540	
ctgagattt	ttctggggaa	gcaggagata	cggttaccctt	acttataatgt	tgtttgttttt	57600	
tggattgttg	aattttggat	ttcttatttgc	aggctttttt	caactgggg	gccttgatcc	57660	
gccttgccttca	gcaatgcac	cgtttcttcc	accgggttcc	tgggaccctt	tcagtgcacta	57720	
tacttagctc	agttcccccac	cctcccaactt	cctaaaaggg	ttaccaggaa	tcctgccttca	57780	
ggtctactgc	cgtttcttgc	gggttgcgtt	agtttcttgc	accaggatgt	aaactcccttgc	57840	
cattccttac	ctgattccag	atctggatc	cgagctttt	accttcttgc	gccaactccctt	57900	
cacttgcat	ttctgtccat	atatcttagt	ccatggagat	acatttttgc	tcttttgcgtt	57960	
tacttacaaa	gtaaaattttt	ctgttttttta	atttttttttt	tgagatgggg	tcttgccttgc	58020	
tcacccaggc	tgtggtgcac	tgacgcattt	tgggtctact	gcaaccccttgc	cctccctgggt	58080	
tcaagcgatt	catctgcctt	agccctccaa	gtagctgttgc	ttacagacat	gacccaccac	58140	
gcccgatctaa	ttttttttttt	tttttagtag	agacagggtt	tcaccatgtt	ggccaggctgt	58200	
gtcttgaattt	cctggaccc	tgatgttgc	atctcggttt	cccaaagtgt	tgagatttttt	58260	
ggcggtggcc	actgtgccttca	gccaattttt	ttttttttttt	attttatttt	tatatgtttt	58320	
gaggataatgt	ttacagtgtt	atatgcattt	ccaaatattttt	gaccaaaaaa	atctccaaaaaa	58380	
aatttagaaat	aaaatccaaa	aaatcttcaaa	aaataccaaaa	aaggcaataat	ctcacagacc	58440	
ataactctgt	accccttcaaa	aaataaaaat	agaaaattttt	cacaactttt	aaaataaaaag	58500	
tacttgcattt	agagaggaaa	gaggaaat	acatccaaat	tacaatgtt	aggccgggttt	58560	
tcacgcctgt	aatcccagca	cttggggagg	ccaaggccgg	cagatccaaat	ggtcagggat	58620	
tcgagaccag	cctggccat	atggtggaaac	cccgtttcc	ctaaaaat	aaaatatttttgc	58680	
caggcatgt	gtatgtgttgc	tgtaatccag	ccacttgggg	ggcttgggg	ggagaatcac	58740	
tgaaccagg	gagacgaaat	ttgcagtgg	ccaaaatcttgc	gccacttgcac	ttcggccttgc	58800	
gtgacaaaggc	gagactccat	cttccaaaaaa	aaaaaaattttt	aaactctttt	gatagaaattt	58860	
ttgggtttttt	tttttggat	gggttgcac	tctgttgc	aggctgggtt	gcagtgggg	58920	
tatgtcagct	caccgcacc	tccatcttgc	gggttgcac	aatttcttgc	tctcggccttgc	58980	
ccaagttagct	aggattacag	gcccacca	ccagaccagg	ctagtttttta	tatttttttttt	59040	
agagatgggtt	tttcaccat	ttggccaggc	ttgttctttttt	ctcctgaccc	caagtgtatcc	59100	

78/122

acctgcttca	gcctccccaaa	gtgctcagat	tacaggcgtg	agccaccgca	ccccacccctag	59160
atagaaaattt	caacatgagg	ccgggcacaa	tggtcacgc	ctgttaatttc	agcacccatcg	59220
gaggctgagg	cgtggggagga	tcacttggc	ccaggagttc	aggaccgca	tgggtgacag	59280
agacagaccc	tgtctctatt	tatttggaaaa	aaaaaaaaaa	aaagagagag	agaaaagaaaat	59340
ttcaacatga	aaagtatctc	tcaaaccctt	cgagatgtt	gcaaaaagcg	actcaaagga	59400
aatgttatta	ctgtgtgtga	atttgcgtt	aaaataagaaa	gaggccgggt	gtgggtggcta	59460
acacctgtaa	tcccacact	ctgggagttc	gaatcaagt	gatcatgagg	tcaggagatc	59520
gagaccatcc	tggctaacat	ggtaaaaaac	tgtctctact	aaaataaca	aaaatttagct	59580
aggcgcgggt	gtctatgcct	gtaaatccc	cactttggg	ggctgaggca	gggtggatcac	59640
ctgaggtcag	gggtttgaga	ccagcctggc	ctacatgtt	aaacctcg	tcttctacaa	59700
atacaaaat	tagctggcg	tttgtgtgg	tgccgttat	ccagctact	cagaggctga	59760
ggcaggagaa	tcgcttgaac	ccgggaggcg	gagggtgcgg	tgagccgaga	tcgcaccact	59820
acactccagc	ctgggcaaca	gcctgggt	cacagtgg	ctccatctca	aaaataaca	59880
aaaattagct	gggtgtgggt	gcctgcgt	gtagtcc	ctacccgg	ggctgaggca	59940
ggagaatgga	gtgaacctgg	gaggaggagc	ttgcagttag	ccgagatccc	accactgcac	60000
tccagcctgg	ggcagacagc	aagactctt	tctcaaaaaa	aagaaaaaaa	aagaaaaaaa	60060
gaaccctgat	aataaaagaaa	ccaaatgtt	aactctca	gctcgac	ttaaaagaaa	60120
taattaataa	aggcagaagt	taaaggggagg	atgataaaag	aattttttt	tttgggtttt	60180
ttgagatgga	gtcttgcct	gtcacc	ttggagtg	gtgatgc	cttggctcac	60240
tgcaacccct	gcctcccggt	ttcaagaa	tctcctgc	cagcctctg	agtagctgt	60300
actacaggct	cgcgcacact	ggcccagct	attttgtat	ttttattaga	gacgggggtt	60360
caccatattt	gttaggtcg	tctcaact	ctgatctc	gtaatctg	cacccctggcc	60420
tctcaaagt	ctgggattac	aggcaggcgc	caccgcgc	ggcttaaa	aaaatattgg	60480
ttctgtgca	aaaggtaa	aaaagagac	atgtttaca	actggagcc	gcacccatc	60540
agctcagtgt	gtctggagaa	aaaacaatct	cgttca	ttcatgat	cgacagccctt	60600
tttgcattct	aaaaatctt	ctatgtgt	tttgaccatt	ctctctctt	ctctctct	60660
tgcttctct	ccagaaaagc	tattcagaca	ttctctctt	tcctcaaa	tccaaacactt	60720
cctccatccat	ccttagcctc	agctgtgc	ctcaacttca	atcattgaga	aaccaggaga	60780
agcatttaag	agtgaacc	cgcctcc	cacggg	accacc	cacaaattt	60840
tgcccccaatt	ctgcgtcc	tccttc	atggatgg	ggtccagg	ccgagccaa	60900
gccaggcctc	ccctggag	ctggatcc	cac	ttctcagg	ggggcccagc	60960
agctccctg	ctccctgt	ccatcaat	ctccc	tgggtc	ccaacaat	61020
atatatattag	tgatgtttt	cccattgt	aaaatactt	agcctct	ctccccccagc	61080
tactatcc	tttgcattt	tccatctt	gcaaaactt	tcaaa	actgttatgt	61140
gctgacttca	ttatcttct	cccggttct	gctgag	tcccac	tctcacc	61200
gttactccat	aaaaatgc	ctgact	ccatcaat	gtgaatgtt	agttttta	61260
tttattcagt	ctttcagcag	catttgcac	ggccgat	tcctctt	aaaaatact	61320
tttctcagcc	aggcgtat	gctcacac	gtatccca	cacttgg	ggccaaggcg	61380
ggaggatcat	gagagcc	gagttcaaga	tcagcgttgg	caacatgg	agaccctatc	61440
tctaaaaaaa	ctaaaaatgt	gcccgttga	ttggatgc	ctgtatccc	atctacttag	61500
gaggctgagg	cagttagat	acttgcgtt	ggggaaatca	ggctgc	agccatgatt	61560
gcaccactgc	actccagct	gagtgcac	gagacc	ctcaaaa	aaaaatagga	61620
aactttctc	agcatattt	tctgat	ctgtgttctt	tgtctgc	gattcgtct	61680
ccttgcgg	ttcttcctca	tccttc	ctttgcac	tgaatgccc	cagagtagac	61740
tctttttt	tttttttgc	acgcgtt	tcgtgt	caagcttgc	tgcataatggcg	61800
aggctcagc	tcatgca	tctgcctt	gggttca	gattctct	cctcagcctc	61860
ccaagtagcc	aggactac	gcacatgcca	ccatgccc	caaattgtt	tatttttagt	61920
agagacaggg	ttttactata	ttggccac	ttgtctca	cttctgc	cgtgaaccac	61980
ccgcctcg	ctcccaaa	gtcgat	caggcat	ccacca	cgccccagag	62040
tacgtctt	agacggc	tctacat	ttgtcccc	tcataaa	cttctgcctc	62100
atggctttaa	ataccatcg	tagactgt	actccat	ttctt	ttttggaga	62160
cgaggtctcg	ctcagtc	caggctgg	tgcgttgg	cgatctcg	tcactgc	62220
ctccacatgc	caagttcaca	ccatctt	acctcage	ctccatgc	tggactaca	62280
ggcacccg	accacgc	gtat	ttgtat	agtagagat	gggtttcacc	62340
atgttagcca	ggatgg	tcgttct	ttttgc	cgcccat	ggccctccaa	62400
agtgcgg	ttatagg	gttttt	tttttt	tttttt	tttttt	62460
cttgctgt	ggggat	tttttt	tttttt	tttttt	tttttt	62520
tatctcaat	gggcaat	tttttt	tttttt	tttttt	tttttt	62580
cctgcagtct	ccaccat	aatgtcc	tttttt	tttttt	tttttt	62640
attcttgc	tttttt	tttttt	tttttt	tttttt	tttttt	62700

79/122

caaatccagt	tagctctcat	catctcccc	gttaccccct	ggtccaggcc	atcttcctct	62760
ctcacctgaa	tcactgcagc	attctccca	ctgggtcttt	tggttctgtt	ttcaactccac	62820
cttagcatag	tctccacaga	gcagtcagag	ggatcctttt	aaagtgtaat	tcccatcctg	62880
tccctgtct	gctcaaaaac	ctgtctgtat	tcccggttta	atctgtcaga	ttaaaagcca	62940
gagctttcc	agtgacctac	atgatctgc	tattatcacc	tcccacttct	ttccccttgc	63000
tcactccact	ccagctctgc	agctgtcctt	tctgtttcct	gaacagccca	gattttgctt	63060
ctttagaacc	tttgttattt	ctgtcccc	tgtctggaaat	gtttttccag	gaagtccacct	63120
ggctctctcc	tgcaattctt	tcctgaccac	catgtttaaa	aatcaatcaa	acacacttca	63180
ggccggacat	ggtggtctac	gcctgtataat	ccagcactt	ggggggccaa	gggtgggtgg	63240
tcacctgagg	tcaggagtc	gagaccaggc	tggccaaacat	ggtgaaactt	cgtctctact	63300
acaataacaa	atagtagcca	ggtgtatgtt	cacacacctg	taatctcagc	tactcaggag	63360
gctgaggcag	gagaatcgct	tgaacccaga	aggcagagga	ggtgcagtga	gccaagatca	63420
cgccacaaca	ccccagctg	ggtgacagag	caagacccca	tctcaaaaaa	aaaaaaagaa	63480
aaaaaaatca	cacaaacaca	tttcccttca	tttccctttt	ccaaagtta	tttttcttca	63540
gaatacttta	cattttttt	atggaaatgc	tccggttccc	cccaactaga	atggatactt	63600
cctgcaggtt	ggcactctag	tcctccatc	caagtaactaa	ccaggtca	ccctgcttag	63660
cttctgagag	caggggagat	caggctgtt	cagggtggta	tggcccagga	atttggattc	63720
tgttttattc	attgctgtt	tgttatttct	cttttgcattt	tcctccctgt	gctgagaaca	63780
ctacttgtac	ataataagca	ttcaataaaat	atttggtaa	tgaatgactt	gttgaatgaa	63840
ttaatctcag	aatatcgagg	ctgggttctac	attagaaaat	ttttcaagg	cattctctgt	63900
tgtcgtaaaca	cattaagaga	ggaaaaatttt	gtactctaaa	tcatttgata	aaatacatac	63960
tgatttctgt	tttcaaaaac	tcttagtggc	tgggcgaggt	ggctcacatc	tataatccca	64020
gcattttggg	aggacgaggt	gggcggatca	cttgagggtca	ggagttgtg	accagcctgg	64080
ccatcatggt	gaaaccctat	ctctactgaa	aatagaaaaaa	ttagccgggt	gtgggtggc	64140
atgcctgttag	tcccagctac	ctggggaggct	gaggcaggag	aatggcttga	acccgggagg	64200
cggagggtgc	agtggccaa	gatcatgca	ttgcactcca	gcctgggtaa	cagagtgaga	64260
ctccatctca	aaagaaaaact	cttagtggat	ttaggaatcc	aaggaagacc	ctcaaaactaa	64320
atagataatc	tagtaccag	aagcccttag	taaaccttaa	cactccatgg	tgaacatcta	64380
gaaacattcc	tactaaaaga	caggcttaaga	atggcttgc	tcttcacccg	tagtccaaga	64440
agtcaaaaag	aagaaatgag	cgctgtat	aaaaataaaa	caaacaaaaa	actaccgtat	64500
cagaggctgg	cagcaaggac	tgaaggactg	tacagtactt	gcctggagca	ggcggatggc	64560
cacaccctctg	cgaagcctgc	tcagctggct	gggggacgct	ccagtgtgt	atggcaggg	64620
tgcagggtac	ttccctctg	agggagttgc	actggggaga	tcctccccc	ctcacactt	64680
ggcagctggg	gctttggat	gtgttgc	ttctgtcaaa	gggtcaatcc	acccctttag	64740
atatgtatca	aaggcgaaca	tatgtatca	aaagggtggaga	acagccaaa	ttaggacttt	64800
taccacatcg	gtggagggtt	acaggcagac	ttggggggcc	tggccagact	tttcatgttc	64860
aaagggtggt	gttggttttc	ctacttctt	tccctccagg	gttccctttt	cctgtgtgt	64920
gaacctgctt	cttttaattt	tttttaactt	ttttaaattt	ttaatttttt	taatttttt	64980
aaattttgaa	aactgtctt	acctgtttt	aaacccctgt	atgatttttt	ttttgttccc	65040
ctgccaaact	gattttttt	cttaatctcc	aaagtggca	tatttttttt	gggttttaag	65100
cagtgactgg	atcatgagag	ctctgacctc	atgatgtttt	taatggat	atgatgtgt	65160
atgggagtgg	catcagtggc	tttataagag	gaagaattaa	gacctggat	agcatggtgc	65220
ccccttcacc	atttgatatac	ttacactg	taggggtctt	gcagaggatc	cccaccaaca	65280
agaaggctct	caccagat	agcttcttca	ccttgcattt	ctcagcttct	gttactgtaa	65340
gaaataatgt	cttttttttt	atgaattttt	catttttttt	tatttttttt	taaacaatag	65400
aaaacgaact	aaggcaaaact	ctcatgtttt	tactgtccat	ccatttcaat	aaactccctt	65460
tatgtttaag	agagccagag	ttggccaggt	gtgggtactt	acgcctgtaa	ttccagact	65520
ttgggaggcc	gaggcagggt	gatcacaagg	tcaggagatc	gagaccatcc	ttgtcaacac	65580
gggtttttttt	gttacttcact	aaaataaaaa	aaaatttttt	tgggggtgt	atgggttgc	65640
tgtacttcc	gttacttcgg	aggcttggat	aggaggagaa	tggcggttgc	ccaggaggcg	65700
gagcttcc	tgatgttcc	tcgttccact	gcactccagc	ctgggttgc	gaatggact	65760
ccgtctcaaa	aaaaaaagaga	gccagatttt	atttctgtt	tttgcacca	agaaatctgg	65820
ctgggtcact	gaagtttca	taaataatag	caattttaaag	actctttca	accaggccaa	65880
tgccttgcct	tgtttagtgc	ttgttgcatt	tttccat	tttccat	tcaaccaact	65940
gtgttccaga	gactaagaat	acaaaaatgg	ggggccgggtt	tgttgcgtca	cacccataat	66000
ccttagcactt	tggggggcc	aggcaggatgt	atcacactgt	gtcaggatgt	cgagaccac	66060
ctggccaaaa	tggtaaacc	cctactctac	taaaaataca	aaaaattttt	tgggggttgg	66120
ggcggacacc	tgtatccca	gctactcg	agactgaggc	aggagaatca	tttgcacccg	66180
ggaggcagag	gttgcagtga	gccgagatcg	caccactgca	cttccagctt	ggcaacaaga	66240
gcgaaactcc	acctcgaaaa	aaaaaaaaaa	aaaaaaagag	ggccggggct	ggcgccagtg	66300

80/122

gctcacgcct	gtaatcccag	caactctggga	ggccaaggca	ggagaattac	gaggtcagca	66360
gatcgagacc	agcctgacca	acatggtcaa	accccatctc	tactaaaaat	acaaaaaatta	66420
tccgggcgtg	gtggcgcaca	cctctagtcc	cagctacttg	ggaggcttag	gcaggagaat	66480
cgcttgaacc	cgggaggcag	aggttgcagt	gagccgaaat	catgcactg	cactccagcc	66540
tgggtgacag	agtggagactc	cgtctcaaaa	aaaaaataaa	aaaaaaaaaa	gaattcaaaa	66600
atttagtagt	tatagtgtgc	ttcttagtta	gttggagagga	catctgtct	tcaaggaagg	66660
ctagaatcta	taccctgagt	ccttactgaa	atcaatccag	cagtcaaaac	atgggaccaa	66720
cgatcacagc	agtaagatg	gaagagcacc	tttgcattac	tagctcatgt	tgagataagc	66780
caactgacaa	gtgtggagaa	gctcacagtt	ctgggttcca	tccttggca	tttaaaaga	66840
aaagtctaa	aaaaattcgg	ttggtcacgg	ttggctcacgc	ctgtaatccc	aacactttga	66900
gaggccaagg	caggcgatc	acggggctag	gagttcgaaa	ccagctggc	caacatggtg	66960
aaaccccgtc	tctactaaaa	acagaaaaat	tagccggca	ttggcgcga	tgccctataat	67020
cccagctact	caggaggctg	aggcaggaga	attgttgc	cccgggggg	ggagggttgca	67080
gcgagtgaga	gcaggccact	gcactccagc	ttggggagaca	gagcaagact	ctgtctcaaa	67140
aaaaaaaaaaag	aaaaaaaaaaag	agaaaggaaa	aaaaagaaaa	aaaaaaaaga	aaaaagaaaa	67200
ttcaggccag	gccaggcctg	gtggctcaca	cctgtatcc	caacactttg	ggaggctgaa	67260
gcgagacggt	gccttagccc	aggagttga	gaccagcctg	agcaacatag	cgagaccctg	67320
tctctataaa	aaaaaattt	ttttggccca	gacgcagttg	ctcagccctg	taatccccagc	67380
actttggag	ggcgaggcag	gtggatcag	aggtcaggag	atggagacca	tcctggctaa	67440
cacggtaaaa	ccccatctct	actaaaaat	acaaaaaatt	aaccggggct	gttggcgggg	67500
gcctgttagtc	ccagctactc	ggggggctga	ggcaggagaaa	ttgcgtgaac	ccggggaggcg	67560
gagcttgcag	tgagccgaga	ttgcgccact	gcactccaga	ctgggagaga	gtgagactcc	67620
gtctcaaaaa	aaaaaaaaaa	aaaaaaaat	taattgtcag	gtgtgttgc	atgcagctgt	67680
agtctctact	actcggggat	ctggggatgg	agatcgctt	gagccaggaa	gttcaaggct	67740
gcagtaatag	tgcctctcac	tctaccctgg	gtgacaatga	gaccctctct	aaaaaagaaa	67800
aaaaaaaggg	aaagaagaaa	agaaagaaaag	aaagagaaga	aaggaaggaa	aaaagaaaaga	67860
aaaagaaaaag	gaaggaagga	agaagaaaaa	aaaagaaaaga	aagaaaagag	agagaagttc	67920
aaagaccaaa	gggtcaggat	ccccaaatag	tttttatgtt	ttattttattt	atttacttat	67980
tttattttga	gacagtatgg	ctctgtcgcc	caggctggag	tgcaatgtat	cgattggcgc	68040
tcactgcagc	ctccaaactg	ggctcagggt	ggccctccac	ctcagccctcc	cgagtagctg	68100
ggaccacagg	cgcggtccac	catggccagc	taatttttt	attcttgtt	gagatgaggt	68160
ctctatatgc	tgcccaggt	ggtctcgagc	tcctgggctt	aagccatcca	cccgccctggg	68220
cctcccaaaag	tgctgggatt	acagaatgt	gccaccgcgc	ctaatccgggt	ggtttgggtt	68280
tttatttgcg	gggtctcgct	gtgtcccagg	ctggagtgg	agtggctgtt	cacagggtca	68340
gtcctggagc	attgcatacg	ctcttgggtt	ctagcgatcc	tccagatgt	ctgcagctgg	68400
gatccaggc	gcccacccgc	ggggggctca	gaatgggttt	ttatattgt	gtttagtgcgt	68460
ccaccttagag	gatataatgt	gtaccgaact	gtgtgcgcag	ggaggcttag	gttgcagtga	68520
gccaagatga	tgccagggca	ctccagcggt	gggtacagag	caagatttca	tctcaaaaaa	68580
aaaaaaaaaa	aaaaaaaat	aaagaattgaa	agtaagggtt	tgaagagata	tttggccttg	68640
tatgttcata	gcagttatgg	ctttgaccca	ctagtcataaa	caaaaaagca	acatgtgtct	68700
gtcagcagg	gaacggataa	acaaaatgt	gtatataatgt	acaattgtat	attattcagc	68760
ctttaaaaag	gaataaaaagg	ctggatgcgg	gggctcacgc	ctgtatccct	aacactttgg	68820
gagactgagg	tggttggatc	acccggaggat	aggagtttg	gaacacccctg	gccaacatgg	68880
tgaaaacttca	tctctactaa	aaatactaaa	attagccgg	catggggcc	tttgcgttgc	68940
atccaagcta	ctggggaggc	taaggcaggaa	gaattgtctt	aactcaggag	ccggagggttg	69000
cagtggatca	agatggcacc	actgcactcc	agcctgggca	acagagttgt	actccatctc	69060
aaaacaaaaca	aacaaaaat	tattattcc	aaagaaaaca	gaccctgggt	ccatttccca	69120
gcccacaccc	gatgttact	cacaacacac	agcctggttt	gtatgtatgc	tttttttattt	69180
aattgtcacc	ttaacttccac	atcccccata	agtcctggaa	taactttttt	ctgaccccttgg	69240
tgtgtgagc	catctccat	tcgtcaacg	tgcaatccct	ctcactgcac	tgagtcaata	69300
gccaacatgt	gtctgactgc	agggtcatcc	ttgggtggctt	aggctgactc	gggcatagca	69360
gggtgctctg	agacctcacc	gcatataggg	tttgcggccca	ataaaactcta	tataatattc	69420
atattatgtt	gtctgggtgt	gtgttagctt	gcactgtctt	ctcgtgacag	tgccctcaac	69480
ctctttccca	ggatttccctc	ctctacccccc	tcaagtccca	ctgctctgc	aagacaaaaaa	69540
gctcgagat	cccagctccc	tcctttacac	cccacgcac	agcctctct	ctcagaacccc	69600
tttaaacaga	gtcttttaat	gcagatccca	agaacagcca	cacccctctc	tcccacccac	69660
tccagacaca	cccaggttaat	tatagcaccc	agggttaacta	tgttagatgga	gtccctggaa	69720
catgtggata	gtgccccctg	ggaggtatgca	aaagcaacat	tgctggcacc	tgcaagagaac	69780
agggtgacat	ccaggaatca	gagcatgggc	ctctggggagg	tagggatgt	gccaggcagg	69840
ctgccaaaaaa	ttggtagagc	aaggccacag	gatctttctg	accttccttc	caaacagagg	69900

81/122

ctcctgtact	ggtgatccct	gtgttgattg	accactccct	tcctgggggt	cgtggtctct	69960
gtcccagttg	cccgacttc	tgtgagtgtc	ctactgaggt	ctttttcatg	agaaggcatgc	70020
tgtccttcca	cctgctggg	gcaagagtga	caacttcaat	actataatag	cagtggcata	70080
cagagaagaa	gaaagatgaa	gtggcaagaa	aaacaggctt	ccaagcagga	gtttttctat	70140
aaaaacaaaaa	acgtttacaa	gcaaaactttt	tataaagggc	tagatagtaa	atatttttagg	70200
ctttagagagc	cacatagact	tgtttgcagg	gactcaatgt	cgctattgt	gtttagaaaggc	70260
agccatcagg	gttatgtaaa	tgagttagtc	tgattttgtt	tcagcaaaat	tttattttacc	70320
aaaacagaca	atgagtggc	tggatttggc	ccatgtatcc	tagttgcca	actctgttt	70380
tggctcacc	catagttgt	tttggattt	ggctctgtca	ctgggtact	gcaggagctt	70440
ggaaggctct	ctgagcttgt	ttccatct	gtaaaattaa	agcaataatt	tctaacaactc	70500
aagagtgtt	cctcacgcct	gtaatcccag	cactttggag	gctgaggcag	gcccgttacc	70560
tgaggtcaga	agttcaagac	cagcgtggcc	aacgtggcaa	aaccctgtct	ctactaaaaaa	70620
atacaaaaag	tagccggca	tgggtggcgg	catctgtat	cccagactt	tgggaggctg	70680
aggcagggtat	actgttagaa	cctggggatg	ggagcgtca	gtgagttggag	atcacacetc	70740
cacactccag	cctggccgac	agagcggagac	tccatctcaa	aaaaaaaaaa	aaaaagagtg	70800
ttagaagggtt	tttagataat	gaataaaaaga	tgccttgcgt	atactaagta	tcaacaact	70860
gatagctgca	ttggtctaat	tataacagtt	tagaagcgat	tgagtcaca	aatgctggat	70920
ttgtcaggga	ggacttctta	tcaggaggtt	gatcttggc	tgagttctgt	agcaaagata	70980
ggcattggat	agaggaggat	agagaaacacc	ctaggactt	tattattat	atgcacacg	71040
gagtcttctt	ctctgtcacc	caggctggag	tgcagttggc	cgatctcgcc	tcactgcaac	71100
cttcgcctcc	cagggtcaag	cgattctct	gcctcttaag	tagctgagac	tacaggtgtg	71160
tgccaccaca	cccgctaat	ttttatattt	tttagtagaga	cagaggttca	ccatgttggc	71220
catgtggc	tcgaacttct	gacttcagg	gatccaccgg	cctcagcctc	ccaaagtgtc	71280
ggaataacag	atgtggcca	ccgcacccag	cccagaacca	tttttcaatc	cttggctctg	71340
ccttttataa	gctgcaagat	ctcaggcaat	tttattttac	tctccaaaga	ctcattttct	71400
cattcacaaa	atgaggcaaa	taataatata	tactatccca	ggttgcgt	agaattaaat	71460
gcaacatgac	atthaatgaa	atgagaagtc	ccttggacat	taactggcta	aagtatgtgc	71520
tgcacaagga	tatcatttt	ggtgatact	tagcatctca	gaactgtatc	tcacaatgg	71580
atatcattga	aacgcattaa	aattttat	aatattttat	aggtatgtg	gcaatttgaaa	71640
gaagaagaca	agaggactga	ttataatgtc	tcaggctcac	tagtctctt	tttaggaggga	71700
aaaacaattt	caagtttaat	tttaggtctt	agattttac	ccctgctgt	cattagaatc	71760
acccagattt	atgaaatcag	agccatctg	aggctgttt	tttcatctcc	agaatgagag	71820
ctgttgggg	gattaagttt	ttgaaaaaaat	acatctaaac	ggtgatcgaa	aatgatagtg	71880
atattattgc	agtgtatgtc	attttttttt	ttattattat	actgaaagag	gttcagttt	71940
tctgtatccat	aaatgtgggg	aattgtatgc	gaccattgt	aagatttctt	ctagctctgt	72000
ttttttttttt	tttgtttttt	gacagagtct	ctgtcgccca	ggctggagtg	caatggcatg	72060
atcttggctc	actgcaacact	ccgcctcccc	ggttcaatg	atccttctgt	ctcagcctcc	72120
gaagtagctg	ggactacagg	cacacaccac	catgcccage	taacttttat	atttttaata	72180
gagggtgggtt	ttcaccat	tgttcagggt	ggtctcaaaac	tcctgaccc	aggtgatcca	72240
cccgccctcg	cctcccaaca	tgtctggatt	acaggcatga	gccactgtgc	ccaaacccctt	72300
ctagctttct	tgtatctgt	ttcttaggtt	ctctgctgaa	atatatttg	gacatcttgg	72360
ataaaagato	atgcaagagc	tcccaatatg	gtattaaataa	ttgattctgg	aggcttagct	72420
actcctgtat	gattagacat	gactcaactg	cctcttctt	gtgtacaaaca	caacaacaca	72480
accaagaaag	gttattctgg	cattccattt	attcagttt	tttacagccc	ttacttccag	72540
cagcacgtta	aaagatatggc	caggccgggg	tgcagttgg	caagtcgt	atcccaggac	72600
tttggggggc	caagggtggc	ggatcacaag	gtcaggagg	tgagaatctg	gcaattcttc	72660
agacttagaa	gcaaccagct	cgataacaca	gttttgcgt	ggctctccct	ctgtccctcc	72720
ctcgcttccc	tcatttctca	tccctggccc	tgagactgtg	cacccatcaca	tagccctgccc	72780
atgagacattt	catctcaggc	tttgcattt	ggggtaatcg	aggctaaaca	ctgagtggcc	72840
ctaaaagagg	attttggat	ggaatgtaga	ttatttcacca	gagaacagac	tttgcgtatg	72900
atcaggccca	gggtgttaatt	tttttttttt	agagaggatg	catagtctt	tctcatctcc	72960
tagtcaaaat	caacaccatg	ataaaataaga	gtcaaaatct	gagatgtgaa	ttggggacat	73020
tttaggtgggtt	aaaccctgaga	agcttgcacc	ttcagacccc	tcaatacccc	tgctccccag	73080
agaaggctgg	acattgaccc	cagcacaggc	aggagccctg	caagatgcca	tttgccttac	73140
taaagatggaa	cccctccact	ctgtttctag	gtaaaataacc	aaagtcaagt	ctccacacag	73200
cctgagcaag	aaagtccag	cctgtctacag	gagaaaaatc	caactggcc	aaaggattca	73260
ctagccctgg	ccactgtgt	tgggaggaac	cagggatca	tgtgtggag	tcaatgttga	73320
agcttggaa	ctgggggtgg	ggtggatata	aagcctggcc	ctggggagtt	tttccctgtt	73380
gagggccctt	acccacaact	caagatccag	tgctatagca	ggagatccca	gagcttagtcc	73440
taacagatgg	tcaggatttga	acttggccct	gagtaaaatg	aggaggatag	tgccagaact	73500

82/122

tctcaacat	actattgagg	aagagggtca	aaggcttaag	gaggtagtgt	aactggaaag	73560
gggtcctgat	ccagacccc	ggagagggtt	cttgaccc	gcataagaaa	gagttcgaga	73620
cgagtccacc	cagtaaagt	aaagcaattt	tattaagaa	gaacacagaaa	aatggctact	73680
ccatagagca	gcgacatgg	ctgtttaact	gagtgttctt	atgatttattt	cttgattctt	73740
tgtcaaaaaa	agggtggatt	atttggagg	tttccaggaa	aggggcagg	atttcccaga	73800
actgtatggat	ccccccactt	ttagaccata	tagataact	tcctgacgtt	gccatggcg	73860
ttgttaactg	tcatggccct	ggagggaaatg	tcttttagca	tgtaaatgtt	ttataatgtt	73920
tataatgagc	agtgaggacg	gccagaggtc	gctttcatca	ccatcttggt	tttgggggtt	73980
tttggccgc	ttctttatca	catctgttgc	tatgacagg	gtctttatgt	cctataactt	74040
cctctgcgca	cctcttat	cctctgttgc	ctaagaatgc	agcctagcag	gtctcagcc	74100
cattttatca	ttggatcgct	ctgattccaa	tgcccttgac	agcaggaatg	ttggaaattga	74160
attactatgc	aagacactgag	aagccattgg	aggacacagc	cttcatttagg	acactggcat	74220
ctgtgacagg	ctgggtggtg	gtaattgtct	gttggccagt	gtggactgtt	ggagatgtca	74280
ctactgtaa	atatgacaaag	gtttcttctt	aaacaggctg	atccgttctt	tattctctaa	74340
ttccaagtag	caccccccgc	ctttttcttc	ctttttcttc	tttctgttattt	tactacatgc	74400
ccaggcatgc	taaggccccca	gctcacatc	ctttccttat	ttaaaaatgg	actgggctg	74460
ggcgcgggtt	ctcatgcctg	taatcccagc	actttggag	gcccaggcg	gcggatcatg	74520
aggtcaggag	atcgagacca	tcctggctaa	cacggtgaaa	ccccgtctct	actaaaaatg	74580
aaaaaacatt	agccaggcgt	ggttgcagg	gcttcgtcgc	ccagggcgtc	aggaggtcg	74640
ggcaggagaa	tggcgtgaac	ctggggaggt	gaggttgaa	tgagccgaga	tttggccact	74700
gcactccagc	ctgggtgaca	gagcggagat	ccgtctcaa	aaaaaaaaaa	aaaaaaaaaa	74760
tagctggca	ttggggcgcg	tcctctgtt	accagctact	ctggaggctg	aggcaagaga	74820
atcgcttga	cccagttaggc	ggaagttgca	gtgagccgag	atcttgacac	tgcaactccag	74880
cctggtgaca	gagtggact	ctgtctcaa	aaaaaaaaaa	aaaaaaaaaa	agacagaaag	74940
aaagagcaca	gacagagtca	caggtatttt	cagtagga	ctgtcagggt	agagtgcacg	75000
gaaataaaaa	gtatatttt	cacttacagc	atcatcttcgt	tttatttagcc	acatttaaaa	75060
tactgaatag	caacgtgtgg	cttatttagt	ttcaactaaa	tcttggacag	tgcaagtcta	75120
agaatccct	gatccgtcc	gcatgggtgc	tcacccctt	atccccagca	ctttggagg	75180
ccaagggtgg	aggatcactt	aaggtcagga	gttcgagacc	agccctggcca	acatggtcaa	75240
acctcgctc	tactataat	acaaaaaaa	tttggccggc	atgggtgtgc	atggctgtaa	75300
tcccaggatc	ttgggaggct	gaggcaggag	aatagtttgc	atccaggagg	cgctgcagt	75360
agccgagat	atggccatgc	actactgcac	tccagcctgg	gcaacacagt	gagactgtct	75420
aaaaaaaaaa	aaaaaaattt	tttggcgtgg	ttggctcacgc	ctgtatctcc	agcactttgg	75480
gaggctgagg	ggggtgatc	acctgggttc	tggagttcga	gaccagcctg	gccaacatgg	75540
tggaaacccc	tctctactaa	aaatacaaaa	attagctggg	ctgtgtgtgt	ggcacactgaa	75600
atctcagcta	ctcaggagggc	tgaggcagga	gaattttcttgc	aaccctaggag	gcagaggttg	75660
cagtggccca	agatccgc	tctgcactt	atccctgggtt	gcagagcaag	actatgtctc	75720
aaaaaaaaaa	aaaaaaatac	tttattgtct	ggacattctg	cagaacatca	tatggagaca	75780
ctatgttgac	gacatcatgc	tgattgtaa	caagaaaatgg	caagtgttcc	agaaaacacag	75840
tcaagacaca	tacatgccag	aaggtgagat	ataaaactcta	ctaagattca	gtggcctgccc	75900
acactggta	cattttaaa	cctgtctact	gtttgttgc	aaaaaggattt	aaccttgc	75960
aaagagggg	ctggccttgc	tccccagact	ctggacataaa	tctttttaaa	ctcttggaaat	76020
atatccctg	atagaagtat	ttttgttttgc	actagggggcc	tttggggccagc	cagatagcaa	76080
caatgtgtac	tgggtgggg	gctttggatc	aggtggcatc	agtgtgaccc	cctgaatggc	76140
tagagactag	aatcaaccac	atgggcagac	aaccctaggctt	acatgtatgg	attccaaataa	76200
agacttttgg	cacaaggggct	tgggttaact	tttccctggtgc	gcaatgtctt	atactgggaa	76260
accacattcg	actccatagg	gagaggacaa	ctggatattt	tcattttggta	cctccctgggg	76320
cttgcctta	tgcatttttc	ccttgcgttgc	tttattttttt	tattatgaga	tggaaatctcg	76380
ctctgtcacc	caggctggag	tgcaatgttgc	tgatctcaac	tcactgtcaac	ctctgcctcc	76440
ccggttcaag	cgattttctt	gtctcggtt	cccggatgtc	ttgggactaca	gatgtatacc	76500
accacacccg	gctaattttt	ttgttatttt	tttttttttttgc	ggttttcacg	tttgcaggaa	76560
ttgtctcgat	ctctgcgtt	catgttccgc	ctgcctcgcc	ctcttcaaaat	gttagaaata	76620
catgtgtgag	ccaccgcgc	cagccccctt	ggctgttattt	tttttttttttgc	ccttgcgttgc	76680
tagtaaaaa	taaccgtgaa	tataacagct	tttagtgcgt	tttgcgttgc	tttgcgttgc	76740
attatcaaac	ctaaggatag	cttggggac	ccctgttactt	gcagtttttttgc	tttgcgttgc	76800
gggtgctcat	gtgtgttacca	tgcctcttca	tttttttttttgc	tttgcgttgc	tttgcgttgc	76860
tttattaccgc	tttactcttca	tgtttatca	cattttatca	tttgcgttgc	tttgcgttgc	76920
cttgcataaa	agacttttca	tctcatgtac	tttttttttttgc	tttgcgttgc	tttgcgttgc	76980
tattttttttt	tttttttttttgc	cttgcacat	acatatttttttgc	tttgcgttgc	tttgcgttgc	77040
tttgcgttgc	tttttttttttgc	gttagtagata	tatgtattttgc	tttgcgttgc	tttgcgttgc	77100

gatacaggca	tgcaatgtga	aataagcaca	tcatggagaa	tggggtatcc	atcctctcaa	77160
gcaatttatac	cttcaagtta	caaacaatcc	aattacactc	tttaagttat	tttaaaatgt	77220
acattaatt	ttgtattgac	tagactca	ctgttgtgc	atcaaataata	atttttttt	77280
tttttagagac	agagtctcac	tcagtgccc	agactgaaag	tgcagtgca	caagctcg	77340
tcacctcaat	ctctgcctcc	ctgggtcaag	cgaatctct	gcctcagcct	cccacatagc	77400
tgggattaca	ggcacacacc	accatgccc	gctaattttt	atattttttt	atgttagagacg	77460
ggtttgcgc	atgttggcca	ggctgtctt	gaactcttg	cctcaaatga	tctgaccacc	77520
tcagcctccc	aaagtcttag	gattacaggc	atgagccacc	acacctggc	aaaatagaat	77580
attcttttagt	gaggctctgt	ggtgacattt	tttttctttt	tttttagagact	gagtctcg	77640
gttgcagct	tgggctggag	tgcaatagca	cgatctcagg	tcactgcaac	ctccaccc	77700
cggatccag	caattctct	gcctcagcct	cccaagtagc	tgagagatta	caggcaccc	77760
ccaccacacg	cggcttaattt	ttgtattttt	atgttagat	ggggttcc	gtgttggcca	77820
ggctgtctc	gaactcttga	cctcagggt	tccacccacc	ttggcc	aaagtgtctgg	77880
gattacaaacg	atgagccacc	acgcacagcc	aatttttcc	tttttgtct	gaatcttat	77940
tttggtcat	ctttgaaata	tattttgtat	ggatataaaa	ttgttgttg	atgttatta	78000
tcattattat	tattttttt	agacagggtc	tcactctgtt	gcctatgt	gggtgttagt	78060
atgtgatctc	gggttactgc	agacttgacc	tcctagggt	caggtgtat	tcccaccc	78120
gcctccctag	tagctggac	tacagatgca	tgccacata	cccaactaa	ttttctattt	78180
tttgtagaga	tgaggcttgc	ccacatttc	caggctggc	tctaaactct	gagctctagc	78240
aatccaccca	ccttggcc	acaaagtgc	gggcatgac	tagccacag	ttacttttta	78300
tagatattt	aatatttaat	atgaatctt	ttgcaccc	tgtaaactgtt	aaaaaaatca	78360
gctgttact	tggcaacttt	ttttttttt	ttttttttt	gacagagtct	tgccctgtcg	78420
cccaggctgg	agtgcagtt	cgtgtatctt	gtctactgca	agctctgc	cccggttca	78480
cgccatttcc	ctgctcage	ctccggagta	gctgggacta	aaggccccc	ccaccacg	78540
cggtgattt	ttttgtattt	ttcgttagat	ttgggtttca	ccgtgttagc	caggatggc	78600
tcgatctct	gacctcgta	tctgtccg	tcggcc	aaagtgtctg	gattataggc	78660
gtgagccacc	gcgc	ttttttttt	ttttttttt	acggagtc	actctgtcat	78720
ctaggttgt	gtacagtgc	gtgtatctc	ctcgtgc	cctccaccc	ctgcctc	78780
ctggccaaata	gctgggat	caggctgca	ccatcacgc	cggtcaattt	ttgtattttc	78840
agtagagat	gggttcc	atgttagaca	ggctggtctc	gaactctgg	ccctcaagtga	78900
tctgcgtcc	ccagcctcc	aaagattaca	ggcatgagcc	accgcaccc	gccaagttagc	78960
actccttta	aggtatct	cttccctac	ccctagca	tttaaacaat	ttttcttcat	79020
tttttatttcc	tgaagtttgc	ttttaataa	tctgtgtgc	gatttttttgc	tatttctttt	79080
gtttcgat	catatgtt	ttttaatttgc	tctgtgttgc	tctgttata	ccacaggaaa	79140
attgtcagcc	tttagcttt	caaataattt	tttgcataat	tctcttttct	cccttttccgg	79200
tacaatttgc	ttttagat	ctaaaaccag	ggccgggtgc	agtactcat	gcctgtat	79260
ccaacacttt	gagaggctga	ggcagggtg	tcaccta	tcaggagttc	aagaccagcc	79320
tggccaat	ggtaaaaccc	cgtctctact	aaaatataca	aaattaccac	gcattgg	79380
acacatttgc	agtccagg	ctgaggcagg	agaattgtt	gaatccagg	gttggagggtt	79440
gcagtgcgt	gagatccac	cactgcgtc	ttggctggc	gacagatgt	gtgagaatc	79500
tgttcgaaa	aaaaaaatgt	tgaatgttt	ataaaactata	tttggtagaa	tttttgttgc	79560
agaataactat	tcattgattt	ttaaacaat	ttagattaa	ccattact	gattttgtat	79620
aattaactt	ctgat	ctca	tgtgtat	aatacaact	gtataaaaaag	79680
actgtgcgt	ggccgggcat	ggtgctccc	gcctataat	ccagactt	ggggggctga	79740
ggcaggcg	tcacctgagg	tcaggat	ttggcc	tgaccaat	gtgaaaccc	79800
catcttact	aaaaaaatca	aattagccgg	tcgttgtgt	gcattgcctgt	aatcccagct	79860
cttcgggagg	ctgtggcagg	agaatca	gaacccgg	gttggaggtt	gcagtgc	79920
gatatcgcgc	cattgcactc	cagcctggc	acaagagcg	aaactccgtc	aaaaaaaaaa	79980
aaagaaaaaa	aacacataaa	acaacaaac	actgtgcac	tttccaaaa	tttaggacat	80040
aattaaagga	actctgtat	aaaattat	ttatcttac	tgtaaaactaa	aatgacttta	80100
tgaagttat	tcagaaat	aatgcagggt	attagttgc	cacagctgc	tattcagcc	80160
aatgtat	tcttgcatt	ttaaatttct	tcttttact	ttactcatat	gtggatcat	80220
aaatttcaaa	agattaaat	acaactct	tagcagca	tttccctaa	catataaaaca	80280
ttttatgg	tgatgattca	gaagggtaccc	gaagaatat	tactcc	tatcattc	80340
ccccatatac	ctgcccgcaca	gacatccat	tttggggacc	tggataat	tgtgggtgg	80400
gagaaagata	ggagaaatgt	gtataagca	atggcttgc	agtctgttgc	acagcgat	80460
aaatccgtc	tctacctt	aacgcctca	tgatcctaca	taagtaccc	cgatcctc	80520
ggccacatct	gtaaatttgg	ggttgcgt	gcagccat	cacagggtct	ctttcg	80580
aaggcagga	attatggatt	aagtgcgt	gtattgtaa	agacttaat	acaaggaggg	80640
cgcataataa	gtacttcata	aataatgac	gccattatca	tgactgaggt	gtatgcag	80700

84/122

gtcggggatt	acggcgactt	cagaatttct	ggtgggagg	gctcaaaggc	agcaaatcac	80760
actggaagtc	gagggtgaggc	actgcttctg	cacagactgc	ttagctggag	agaatgagga	80820
aggcttagag	gagattttaga	ggaactttaga	gtccctccccc	tccaactctg	tgggatctgc	80880
tccctgtccca	gagacatca	ggggatttct	gtcactctcc	cctcccttac	gtccctccccc	80940
ccccatccaa	ctaaccacac	aacacataca	aaatagcccc	tgcgaggttc	tgcacgctgg	81000
aagggAACAG	gagaaggggcg	ctgcgccttc	ttgctgtatgc	cctgtacttg	ggcccctgg	81060
agacacAGCC	acttgtcccc	tcagcctgca	gagaatccc	acgttagacccg	cgccccgggtc	81120
cttggcttca	gccaatctcc	ctttttgtggg	gggtggatgc	acgtatccaag	gttttattgg	81180
ctacagacag	cgggggtgtgg	tcggccaaga	acacatgg	gtctcccgagg	gcatactcgga	81240
tccctgtgg	ggcgcgcgtc	agcctccccc	tgccaggccc	gccgaggccca	ggaggaagcg	81300
gccccACCGC	gtccattcgg	cgccagctca	ctccggacgt	ccggagccctc	tgccagcgct	81360
gctccgtcc	agtgcgcctg	gacgcgtgt	ccttaactgg	agaaaggctt	caccttgaaa	81420
tccaggcttc	atcccttagt	agegtgtgac	cttgagca	tgacttttatt	tttcagtgc	81480
tagtttcca	gataccagg	ctgactccaa	ggacttac	tcatctggag	gttttagcac	81540
agtaccgtcg	catatgtat	ttccatgtca	ttttttgtt	ccttttcatgc	acttgcaaac	81600
atggcatgt	ctgaaacaa	ataggcacat	ttttttttt	tttttttttta	aggagtcttc	81660
ctctcgccca	ggctggatgt	cagtggcg	atcttggctc	actgcaaccc	ccacctcccg	81720
tgttcgagat	tctcctgcct	cagccctctg	attagctgg	actacaggca	tgccacgacg	81780
ccccagttat	ttttgttattt	tttagtagaga	cggggtttcg	ccatcttggc	caggctggc	81840
taactcttga	cctcagggtg	tctgactgc	tcagccttc	aaagtgttgg	gattacaggc	81900
ataagccact	gcatactggcc	agaaatgaaa	taagtaatc	tttaaacctg	ctctaacaat	81960
atagtaaaaa	gaccatatta	ttattagagc	aggtaaggg	atttgcttat	ttcgggttct	82020
agttatagtc	ttaaaacttgg	acattttgt	agaaaagtaa	aagttttctc	ttcaaaggttc	82080
cccttcttgc	taaaaagatac	atcataatgt	tttagaagtaa	tagtttattt	taaagactaa	82140
ctttcttca	gcctccttgc	ttttgtctaa	taactttttt	ttaagcccta	tccttatgtaa	82200
cttgggaca	tgctcacagg	cacgttccag	ttcacagctc	atgccttctc	tttatttgg	82260
aatgttattt	cttcctttaaa	ccttcggta	agcaacttcc	tctccttctt	cgtttcttc	82320
tgcacttacc	tatttagaaa	gttttaggtc	attagcaaat	cggtctatcag	ttaagagtg	82380
tgaggtcccg	ctccagccaa	tggatgcagg	acatagca	gaggacgacc	caaatgcgt	82440
aggataaaat	atgtttgtt	ttccctttgtt	cagggtgt	ctcgacatcg	tccatctgc	82500
gattgagcac	ccttctgc	gaaaatggaa	attgccttgc	tggagatctt	ttgtctccgt	82560
gctgactttt	cttcgtggca	ccgattatct	atttctaaca	attttggat	ttctaacatt	82620
ctgaacaatc	ttgggctagt	tgtcttctt	gggcctgtt	ccccatccgt	cacatgataa	82680
acttcattgg	tttaaaaacc	ccagcgaaca	tttatttgcgt	tacttattac	ttctctggcc	82740
ccccaaaccc	aaccccaagg	agcgttaca	acctcagccg	ctgagcgcac	tcgccccgtg	82800
ttaagaagca	ccaaagacag	ggaggcttg	ttgattttc	tttggagta	gagggtcaga	82860
agattcacag	gaaaatggca	tttgagcaag	gatgattcac	tggagctagc	ttttaaataac	82920
tggcgaggct	tttatgttgc	agtccttac	aaagttgagc	attcgacagg	actgcactcc	82980
gaaataagcc	cgcttccct	tttatttgc	taatgatcca	gggagctgt	gttccgcac	83040
gcccgcagg	gtgccttttc	ctaatcagg	ttctgcatcg	cctcgacaaa	gcaggccgt	83100
gccccgttcc	ctgagga	aggactggg	gtgcagggt	aaagtctcg	tgccggccag	83160
cgcctgttag	caaaactcaa	acggaggagc	aggaggggtc	gagctggagc	gtggcagggt	83220
tgaccctgca	ttttagaaagg	gcacaattt	aaagggtaccc	aggggcccga	agccggggac	83280
ctaaggcccc	ccccgttca	gtgtctggg	gggctccccc	cccagggggt	tagtttgc	83340
gagactgggt	ctgcagcgct	ccacccgggg	ccggcgcacag	acgcacaaa	acagctgc	83400
gaacgggtggc	tgcgttccagg	caccagg	ccggggaaa	ggcgcgggt	gcacgcgcgg	83460
gtcaacgtgg	cgatgcggc	gtgcgttcc	gcaccccg	gagggggatg	ggggaaaagg	83520
gccccggccgg	cgcttgcac	cccgta	ctagcgcgg	gaaggaccgg	aactccgggc	83580
gggggggttg	ttgataat	ggggctgga	gtgcgttgg	catcccgagg	aggcgggtgg	83640
gccccactccc	ggaagaagg	tccctttcg	cgctgtgt	gccccccctc	ttgacccgg	83700
agtccggggc	ggttgtctaa	tgaggggg	ccggcccttc	ccgcgcgggt	ccccccgcac	83760
ccttcgtccc	gaccggggcc	ccggcatgtc	tttcttccgg	cgaaaggta	gtcgaggggg	83820
cgccggccggg	gagtca	gggcctcagg	ggcggcgggt	gggcagggt	gcctgcgagg	83880
gcttcccca	aggcggcagc	aaggcctca	gchgagcctcg	acctcggcgc	agatgcccc	83940
tgagtgcctt	gctctgtcc	gggacttcc	ttggagggag	aagggtggct	tcttgcgcga	84000
ggtcagagga	gtattgtcg	gtcggttc	aaggcattgc	taaagccat	agaagttcc	84060
gcctgtttgg	ttaagaacag	ttctttaggt	ggggtaggt	ttttttgtt	tctttgagga	84120
ccgtggatca	agatcaagga	aatcttta	gaaccttatt	atggaaagtct	gaagtttcca	84180
aatgttggagg	gttttatgtc	taaaagcaac	acgtaaaaaa	attgtttct	tcacccagtg	84240
ctgtcttcca	atttccttctt	ttgggggagg	ggtagttact	gctgttacta	aaataaaaatt	84300

85/122

acttattgct	aaagttcccc	aacaggaaga	ccactacttt	tgatgacttt	ggcaagtttg	84360
ctaactactg	gaaccctaac	ttacaaaacg	actacttaca	tttttgattt	ccagtgttat	84420
tacctgccc	atgttaactg	agaaacagct	taattttgat	tctgggttac	gttggtcac	84480
ttcattaaaa	atacatatcc	gaagtggagc	agtatgggtc	tgtggacagc	agtgattttt	84540
cctgtcaatt	cctgttgc	cagataaaat	gtaccagaca	gaggccggc	gcccgtggctc	84600
acgcctgtaa	tcccagca	ttggggaggct	ttgggggtgg	atcacctgag	atcggtggatt	84660
caagaccagc	ctgaccaaca	tgggaaaacc	ccgtgtctac	taaaaataca	aaatttagcca	84720
gggtgggtgc	gcatggcgt	aatggccatg	acttggggagg	ctgaaagcagg	agaatcgctt	84780
gaacctggga	ggcggggat	gcccggggat	gagatagcac	cattgcactc	cagcctgggc	84840
aaaaagagcg	aaactccgtc	tcaaaaaaaaa	agtaccagac	agaaaatgggt	tttggtttct	84900
tttttggtt	tgagacggag	tttcgctctt	tttgcggagg	ctcgagtgca	atggcgcgtat	84960
ctcagtctcg	gctcaactgca	acctctgtct	cccagggtta	atcgatctc	ctgcctcagc	85020
ctcccaagta	gctgggat	cccatggccc	accatggccg	gctaattttt	gtatttttag	85080
tagaaacggg	gcttcacat	gttaggtcg	tcttgaaccc	ctgaccctaa	gtgggcctcc	85140
cacccctggcc	tcccaaagt	ccaggattac	aggcatgagc	caccgcggcc	agccagaaaat	85200
gggttttgg	aaaagacta	aacaaaatcg	aacttggttt	catatgacag	ctctgctgct	85260
aactgttaaca	ggggcagacc	agttaaaccta	cttttctgtc	ttctgtcagc	tgagaatttag	85320
atgattccca	aaggccatt	gaactctgaa	tgactttttaa	tacttttct	taatgtggta	85380
cacccgtttt	gtaaactgat	ccaggatgt	atgtcatgaa	agtgttaat	gaatgaaaacc	85440
ggtaaaaatag	taggaggaag	ctttatttgg	aaggcagggg	tatacttaat	agctctctaa	85500
tttatttggta	ttgaagtgg	taacttttg	tttttaagg	ggggaaaaca	ttttaagaat	85560
aatgaggcaaa	actgcatatt	gcacaagaga	ctgttgtctc	tattcaacaa	ataccctttt	85620
agtgtccaga	gtctgccc	tgctgtgtca	ggcccttcac	attgagtagt	gaaccagaga	85680
atgtccctgc	accatggag	cttattgtct	acttgggtt	acagataata	aataagcaaa	85740
caaattctct	ctcttctccc	tttcgctcc	tgttaatgt	tgtgtatagg	tgatacttta	85800
caagttgagt	aaagtgttat	gaaagattaa	gaggagaaaat	gcattttgt	tagatgttag	85860
aggactcage	aggtagcatt	gaaacttaga	gctgaaggat	cagtagggagg	taacttagaga	85920
ggccaggggaa	tcgcatgtt	aaaggccagg	aggcagaagaaa	gagcatgttgc	cccttcaaga	85980
gaggaaagaa	ggctactgt	actggagcat	agatgttaggc	aagtgttgg	tgatttagag	86040
cttacggggc	catgtttag	ttttatttcc	atgtccgaga	tgccaaacat	gttgggttcat	86100
atctgttaatc	ccagttat	aggaggccga	ggcaggaata	tagctgaac	ccaggagttc	86160
aagaccagcc	tgagcaacat	gagacctgt	caaaacattt	aaaaaattgc	tgggtatgtat	86220
gggtcacacc	tgtgttccca	gctactcagg	aggctgagc	agaagatca	cttgcgccta	86280
ggagggtggag	gctacaatga	gcccattttt	agtacttaca	ctccagctc	gtgacaaaag	86340
tgagaccatg	tgtcaacaa	aatacagaaa	aatattttat	ttttttttttt	gaaagaggag	86400
tgatctgaac	tttatactt	aaaagatcat	tctagggcat	ggggctcat	gcctgtatc	86460
aagggttttgc	ggaggctgag	acaggaggat	cacctgaggc	cagttcgaga	tcaacctgt	86520
cagcatagag	agactccatc	tctacaaaaa	gaaaaaattaa	atagctgggt	tttggtaggt	86580
attcaggagg	ctgaagcaga	aaatcactt	gagccccagg	ttttggggat	gtactttttt	86640
atgatccccac	cactgcaaca	cagttagatc	ttgtctcaaa	aaaaaaaattt	aatcatttct	86700
gggtcttttt	ggaggctgag	tgtgtttaaga	gtagaagctg	gagatggttc	ttttttttttt	86760
tcgattcaga	ctttaaatac	catcaatgca	tttgggttcc	tttgggttcc	ttttttttttt	86820
atccctggcc	ctgaatccat	actgttat	ttttttttttt	ttttttttttt	ttttttttttt	86880
tgaccaat	agaggtgtc	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	86940
tcttgagcca	cacatagat	actacaacgc	taacaatagc	ttttttttttt	ttttttttttt	87000
gcaaaactta	taatgtttt	agaaaagttt	cgaattttgt	ttttttttttt	ttttttttttt	87060
cctggccgc	gggatggaca	agcttaatcc	ttttttttttt	ttttttttttt	ttttttttttt	87120
attttatgcc	agattttagt	atttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	87180
tttgggtttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	87240
agtgcagtgg	cggtatctgg	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	87300
tgccctcagcc	tcgcaagtag	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	87360
tggagacagg	gtttcaccat	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	87420
tgccctgcctc	ggcctcccaa	aggctgggt	ttttttttttt	ttttttttttt	ttttttttttt	87480
ttttttggatt	tttagttgt	cagccaaaaa	ttttttttttt	ttttttttttt	ttttttttttt	87540
tcctactcta	tatctgtatc	atcagcaat	ttttttttttt	ttttttttttt	ttttttttttt	87600
cctaccacgt	ctcaccat	gtgacatta	ttttttttttt	ttttttttttt	ttttttttttt	87660
agattgagtg	gttaaggatg	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	87720
aagagggagc	ttgttttata	aagattgagg	ttttttttttt	ttttttttttt	ttttttttttt	87780
ttgggtttcca	ttccattttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	87840
gaccggat	tattggaaat	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	87900

86/122

ctgtaaatga	tgctagttat	ccatcatctg	gaaaaaaaata	atttcctgcc	tcctctcata	87960
tatctcgat	caacagactt	tttctgttaa	gggc当地atc	ataaaatattt	taggctttcc	88020
agaccatgt	gtttctgtca	cacttcctt	tatccttgaa	gccc当地agaca	atatgtaaaac	88080
aaatggcgt	ggctgtgtca	cgataaaact	ttacttacaa	aaaactgttag	tgggccagg	88140
taggcatggc	cagcaactt	ggaggctaaag	gcagatggat	cacttgggt	caggagttt	88200
agaccagct	ggccaacatg	gtgaaaccct	gtctctacta	aaaatacaaa	aatatagctgg	88260
gcatgggtgt	gggtgtctat	aattccagct	actctggagg	ctaagacaca	agaatcactt	88320
gaaccaggaa	ggcagagggt	gcagtgcgt	gagatagcac	cactgcactc	cagccagggt	88380
gacggagtct	taaagaaaa	caaaacaaaa	gttagtgggt	tgtatggc	ccatgggctg	88440
tagtttgc	atccccatg	cagaacaaaa	tccaggtaa	ataagagct	ggaatgttaa	88500
aaaaacaaaa	cttgaagtc	tgtagaagaa	caggtagggg	gaacaatct	gatctcagga	88560
tagaaggga	tattgtttaa	aataagacac	aggaaaatat	aatccatgtt	gtgtaaaattt	88620
gactacgtt	aaacaaaa	cttccggcaa	ggccgggtggc	tcacgcctgt	aataccagta	88680
cttgggagg	ccgggttgg	cagatccac	gttcaggaga	tttaggaccat	cctggctaac	88740
acggtgaaac	ccccgtctc	ctaaaaatac	aaaaacattag	ccggccgtgg	tggccggcgc	88800
ctgtactccc	agctacttgg	gaggctgagg	caggagaatg	gcctgaaccc	gggaggcga	88860
gcttcagtg	agctgagatc	gcccactgc	actccagct	gggc当地acaga	gtgagattcc	88920
gtctcaaaaa	aacaaaacaa	aacaaagcaa	aaaacctaaa	actttcatac	aataaaagtat	88980
acctaagata	tttcttagaa	agaagattt	cattccaggac	gtgtatggaa	tttctgcgaag	89040
taataagtaa	aagacaaggg	acatagaaag	gcagttcaca	aaagaggaaag	ccaaaatgtac	89100
caataaaacat	gaaaaggat	ttaacactaa	agggaaacaag	gaaatgaaatt	aaaaacatca	89160
aatgcattt	caaaaactagt	aagtggca	aattaaaaat	accaaggatg	agaatatgaa	89220
gcatggctat	atgagtgc	ggaatggta	agtcaactt	attaaaaatg	cacataattt	89280
gttttttatt	tatTTTTT	agacagtcta	tgtcgccca	gcttagatgc	agtggcatga	89340
tctcgctca	ccacaatctc	tgccttctgg	gttcaagcaa	ttctctgccc	tcaagcttct	89400
gagtagctgg	gattacaggc	acatgcacca	acgcccgggt	aagtttgtt	tttttagtag	89460
agacagggtt	ttgcccattt	ggccaggctg	gtctcgaact	cctgacctca	ggtgagctgc	89520
ttcccaaagt	gctgggatta	gaggcgtgag	ccaatgttcc	tggctgaaaa	aatgcacat	89580
aatttgttac	cttagcaattc	catgttca	ggcttatcc	agagaaaaat	ttctgttat	89640
gcattagaaag	acgtgtacta	gaatgttac	tagttgaat	ttaatgtt	atttaggaaa	89700
taaagtaaat	gttcattaac	aggaaaatca	gttaaggat	atttataaaa	caattaagta	89760
gctaaaatga	ataaaactaga	gctcgtgaa	tgaacttagaa	ctggttcaat	agtcatgtca	89820
gattattgaa	tgaatacagg	tcagatatgt	atagagtgtc	atttgtgtt	ttaatttttt	89880
ttttttttt	gagatggagt	ctcactctgt	tgcccaggct	ggagtgcagt	ggcgtgatct	89940
cagetcactg	caaccccccac	ctcctgggtt	aaagtgttcc	tcctgctca	gcctcccgag	90000
tagttggat	tacaggcat	caccatcat	ccaggttcat	ttcccttattt	ttagtggcca	90060
cagggttca	ccatgttggc	caggctggc	ttgaactctt	gacctcaagt	gttccacccca	90120
acttggcctc	ccaaagtgt	aggattacag	gcgtgagcca	ccgtgcttag	coatttgcgt	90180
gattttaaa	gatgtgcaga	ataatgcat	taaaaaaaaat	acacatacat	gtatatatat	90240
acacgtttgg	ctgggttgg	tggctcacac	ctgtatccc	agcactttgg	gaggctgagg	90300
caggaggatc	acttgaggcc	aggtgtacaa	gactagctt	ggc当地atgat	caagacccca	90360
tctcaacaac	agaaaaggata	attaggtatg	gtggcatgag	aggatcaactt	gagcccaggaa	90420
gttcgagtgt	tatcaggcca	ctgcaactct	gcctggacaa	caaagcaaga	ccgtgctca	90480
aaaaaaaataa	aataaaaaat	atttggatgt	ggtcatagtg	aaaaaaacgt	catggaaaggaa	90540
aatatgtctt	atttattttat	ttatTTTTT	tttttaaga	cagatcttgc	ctctgttacc	90600
caggctgggg	tacagtgggt	taatctcagc	tcaccgcata	ctcgccctcc	cggttcaag	90660
cgattttct	gcctcagcc	tctaagtagc	tgggactaca	gttaccgc	accacacccct	90720
gctaattttt	gttggggat	tagagacagg	gtttcaccat	gttggcaagg	ctggcttcteg	90780
acttctgacc	ttaagtgtac	caccgcctt	ggcctcccaa	agtcctggga	ttacagggt	90840
gagccactgc	gcttggccag	gaaatatcta	attttagta	tgggaaaggaa	90900	
agggtcagggt	ggtgatttc	aggaatctt	aatcttgc	gggggacacca	90960	
agaaaataaa	gcaaaatgt	gatatttgc	tgttgcgtt	tgtatatgtt	agaagttataa	91020
cataggagat	ctgattgtata	gttaggagaat	gttttttaggt	gttaaaatgt	gaaccgtgt	91080
ggtttgggg	ggcagtagaa	tcagggttgc	atagttgtt	tgtgaaagg	aataaaacaga	91140
ccatgttaag	gatgacttcc	ggaattttgg	tctgttgc	gggtggatgt	cagtgttacc	91200
cataggggaa	gatgaagat	gaggtaggaa	cagggttggg	agaagatgtac	atgttccctt	91260
ttagacaagt	gaaattatgg	aagatggcag	gttaggttgc	agctatatgt	atttgagata	91320
aaagattttag	gatggagata	taaattttagg	agtaacagcg	tatctatgtt	attgttgc	91380
ttaagaatgg	gtaggatcag	ccagaaaata	cagatgtata	tgcagaagag	aggagtcaag	91440
gaagccaaga	caagttatg	ttttaaagtgt	gtgtatgttgc	ccatggccag	atgttgc	91500

87/122

gagggtcgca	aacaccagt	accctacaac	attttaaat	gtcgtttcc	tgacagcagt	91560
gatcagtacc	tgcaacgatc	ttatttattt	tttcatgtt	agtctccaca	cacttgaatg	91620
tagactttt	gaaggcaaaa	tcattgcctt	ttctgagctg	ggagcatgtc	tggcacatac	91680
caagactca	acagttgatg	tattgacttc	atccagatac	tctgaggcg	agttatttcc	91740
tgctactagc	cttgcacctt	tcaatgttta	agagcacaaa	tacagagatg	ggcacgttt	91800
ggcatttctt	attttgataa	cctttccctg	gtaaagattt	ttaatgttga	aaaaaaaaaa	91860
caagaaaaga	gggttaaaaa	tagtcttgc	ttagatccctg	tgatagaatt	cacacttggc	91920
ttaactgtgt	gggcacccctc	ctatcttgg	tgtcatattt	gcttatctac	agcagaattt	91980
ttactgtttt	atgttagtaag	gaagcaatta	tatgattttt	ttacagacaa	attattcttt	92040
atctttttt	tttttagacg	gagtctctt	ttgtctccca	ggctggagta	cagtgtcgcg	92100
atctcggttc	actgcaacct	ccgcctccctg	ggtcaagca	attctctgcc	tcagectccc	92160
aagttagctgg	gcttacaggt	gtccgcacc	acacccagct	cattgttttg	tattttttagt	92220
agagatgggg	tttcacccatg	ttggccaggc	ttgtcttgc	ctactgacct	cagggtatcc	92280
acccgccttg	gcateccccaa	gtgctggat	tacaggctgt	agccacccgt	cctggccccc	92340
acaatttttatt	atactctgag	tttttagaggc	tttagatgtt	ttcacttgat	gctatgggag	92400
gaataaagtaa	taagatatga	tacacaacca	aagacccccc	ttcactatgc	ttctagtagc	92460
tagtactatg	gatgacacat	ggtaataata	ttggtttagca	tttgcctc	atttactgtg	92520
ctagttactc	ttctaaagcc	cttacaggta	tatatttttt	ttcatcaata	atccctctaag	92580
gtatgttttta	ttatttgactt	atattttataa	atcaagaaaa	ttaagaccca	gagaagtaag	92640
taacttgtcc	aaagatcacat	ggcttataag	ttgttagagcc	agaatttgac	cccagatgtt	92700
gtgactacat	tgtctctcca	taagcaggtt	caactctttt	gactggatgc	tgttccaagg	92760
tcacttcctt	agagaagctt	ttgctgacaa	ctaccctctt	gtgcctccct	ccaaggctgt	92820
ccattgttct	agaactttga	atactccatct	tagataaaag	ctggcttaat	tttacagtg	92880
ttatagaaatg	gatctctgac	tgcaaaaatg	ggtcataatt	atctttttat	gttctagtg	92940
aaggcaaaga	acaagagaag	acccatcgatg	tgaagtccat	taaaggttaag	ttctggccctt	93000
ggcagtccac	tgcataaaaa	agtgtatgtc	tttgcatttg	tgagttctt	aatctgttta	93060
tactctctct	tttggcattt	atcatttctg	ccttatttttta	taattactta	tgatttttgt	93120
ttatttccct	ctttaacctg	tataatgtt	taacatctag	catataataa	gtaggcttt	93180
ttttttttt	tttttttgg	gacggagtct	ttgtctgtt	cccgaggctt	agtgcagtgg	93240
cgcgatctt	gtcactgca	agctctgtt	cccggttca	caccatctc	ctgcctcage	93300
ctccccagca	gtctggacta	cagggtcacg	gcgcacgc	tggctaattt	tttgcattttt	93360
tttagtagaga	cagagtttca	ccatgttagc	cagtatggc	tcgatcttct	gacccctgt	93420
tccggccgc	tcggcctccc	aaagtgtgg	gattacaagg	gtgaggccacc	gcacccggcc	93480
gtaagtaggc	tttttttgg	ttaattttat	tttttttgg	tggaggctt	ctcttatccc	93540
caggctggag	tgcagtgg	ccatctcg	tcactgcac	atccacacttcc	cggttcaag	93600
cgattctct	gcctcagcct	cccggatgc	ttgggattaca	ggtggccgg	accatgccc	93660
gctaattttt	gtattttttg	tagagacagg	gtttcacccgt	gttggccagg	ccagtctcaa	93720
actctctgacc	tcaagtgtat	cactcgctt	ggcctcccaa	agtccctgg	ttacaggcgt	93780
gaggccatcat	gcttggccat	aaggtagctt	ttactgcac	ttgtgttgc	tggtatccct	93840
agtatttaca	gtgaaccagt	gcccttctt	ttaatcacac	attnattgt	tccctaaaag	93900
tgatttagttc	atcttttttt	tttagtaaga	caaaaaatgt	agaataactct	taactgagca	93960
gtctgttaac	tgttagaaag	cactgacact	tataaggctt	agtttttgc	catttatcca	94020
gaagtatgtt	tgatttacat	tttttttttt	ttatgttgc	gaacaaccc	attnaaaaat	94080
atattttgtt	tatttttttt	ttggatcgat	acattgtctt	tgtttataga	ttagagcatg	94140
ctttttaaag	atgtcttatt	actactgtat	tttattttgtc	cagtgtacag	agattgaagt	94200
ggggaaaattt	taatggaaat	ttgttccata	gtcattacat	attaatttca	tcaattttatt	94260
tccataaaaat	ctgttagattt	ctacttattt	agattttttcc	ttcaaatgtt	tttatgttgc	94320
attgtctgca	ctggatgttt	attctatatgt	tcacatgttgc	tggagaagaa	gactaatttat	94380
aacttaggca	agtgtaaaaa	tttagggaaaa	agtaaggtt	ctttacagcc	tagttactt	94440
atttctttagt	taaaggccat	tagattccac	attagttcaa	actgccttct	tttagccaaaa	94500
cttgatttggc	agtgtatcaa	gctttaagcc	tttctcaagc	agagacctgt	aaagactaga	94560
tctgactgtt	gtagaagaa	ggaactttaga	tttttcaggc	agtggagaaca	ccagtcttcc	94620
actctaaact	ttggccatca	cagtatgacc	ttgggaaatgt	gtacttttct	tcagatttctt	94680
catttggta	atggggggat	ttggcttagt	atttttcaaa	tctctactgg	gtcaaaaaat	94740
tctgtcttta	tactctgtt	atgttagtaca	taatctgtgc	ttaacattca	ctgacttattc	94800
cttaggatata	tacagaagca	gtacaagaaa	cagccctca	agatgtttgc	agtctgggtt	94860
gaaagacaaa	tttatacaca	gaacagttagc	aatatagacca	aaataataat	actgcccatt	94920
tatagaacac	ttttctctt	ctggcattt	gacaaaaaaat	gactataacg	gtgaacaaaaaa	94980
aagactttagg	teetggccctc	attgaactt	cagatttagt	ggggagagga	acattaatca	95040
agtaatttcca	agatggctt	agccatgatt	ggttagtgc	gaagtaaaga	gatgtgaacg	95100

gacttggaaa	aaaattcggg	ggcaaaaatgg	atagaagttt	attattgatt	aaatatgagg	951
tgtgagagag	agggatattt	aagattgata	cctacccct	ggcttcgccta	acagaaccaa	952
aacaggaat	tatatgttca	gtttgttat	gttgggtggg	agggtctttt	gagtcatca	952
tttatatatg	ttatatatgt	tatTTtat	gcatagtaat	ttaagggtct	gagttttaaa	953
ccaaagggtt	gagagtgatt	tttagagtc	tagcaaccc	aagtggaaaat	cctgcctgtt	954
gaaatggctg	tttactagct	cattaaccc	gggcaaaatg	ttaacttgtt	tttcattttt	954
gtcttcatct	ctaaaaatgg	gaaaatgg	tcttacaaga	ttgtcctgag	agatagatga	955
aataatatcc	aaaaaaaaaa	aaggtaata	gagaactcg	tatagtgcct	ggtatatagt	955
aggtccttca	ttggtagcta	tcattatcta	gttttaacat	agcccttcagt	ttgttgatt	956
agtcaactg	agtgaagcac	tgcaaggaaat	tcagaggaat	ttgagatcaa	caaatgattt	957
ctgaagttt	ggggagactt	catggcaat	acacttacct	tgataaaaag	ttgaagaata	957
agaaaagttt	gaatggagaa	ttctttctct	tctccctacc	agcccagctt	tttatttgag	958
gatataattgg	gcaagggggc	cttcagacaa	gtagagggag	attttacag	aaagatttgag	958
atgaaggat	agaaggctgt	aaagaccaga	aaagagaatt	gagacagagg	aagcaggaaag	959
ccactgttag	tttttgagca	agatattgtat	gctgttaagta	tgtgttttat	gaaagggttag	960
tctggaaag	atttgcggaa	tggagacccc	ggaaggtttt	ttgttataat	acagaaagac	960
ttgcaactg	ggtgagggt	aaaaataaa	caggttaagta	aatgtttaaa	catcttgaag	961
gaaaagtca	caaatcttgg	caagtaaca	gataacagtg	aaaagaatg	ggaccaagat	961
tttgagttt	ggagactgg	ggattgaaca	gacaggaaa	ttgagaggag	aatcagatga	962
tgatgtttt	agttgtat	tagacagat	gtgttgaga	tggttaaagt	aatgtgggtg	963
ggaatgttta	gtgcgagta	atcgtgata	caagacaaa	gcccaggctca	aagacaagtc	963
acagatacag	atcagggttt	tttcatctgc	tccacagagg	tgtaccctag	gagctgttgc	964
aaacagtcca	tgtggagggt	gtgagtaaga	tgtttccctt	gaatttgccca	gaatttactt	964
tttgggttg	ttttgtttt	ttctgagaca	gatttcgtct	ctgttgcctt	ggctggagggg	965
cagtggcgag	atcgcgcage	tcactgcaac	ctctgcctt	cggttgcag	tgatttctt	966
gcctcagctt	ccccaaatgc	tggattaca	ggcttgcctt	accaagccca	gctaatttct	966
tttgtatTTT	tagtagat	ggggtttccac	catgttggcc	agactggct	cgaactctg	967
gcctcgat	ctgcctgcct	cagcctccaa	aagtcttggg	attacaggcg	tgaaccactg	967
cacccgggcc	cttggtaagt	ttatTTTgt	gggaagcaaa	ggaggtttca	gtttttaaaa	968
agtttggaaa	ttatgtctt	ggtaatTTA	aaagatttga	gagtaataat	gttttcttagc	969
agaaaagata	aaagaagaac	agatagcttc	aagaagggg	gccaagaag	caggctatat	969
ctgacacact	gggtgttgc	aatatggat	taaaagaatg	agagcaatga	gcagatagaa	970
gagggaaattt	ggagagtata	atccatgg	gaccaagaaa	gatagactat	caggaaggag	970
tggtaaaaaat	aagttaactg	ttctaagaga	gatgttaga	gggaccgggg	aaagccttgc	971
acaatagt	tagtagattt	ttacattata	catatcta	taagaaacaa	tgcgagatgc	972
tcaccatttcc	tatagactt	tacttgact	tgtctgaaca	cgaaaaactgg	cttttggttt	972
taaataaagct	aaaaattatt	ttgctccat	ttctcatgg	taaaaaaaa	aaccttcttt	973
taacattgaa	aaaatagttt	gaagacatgc	acttttcatt	ttgttaattc	cacaactatt	973
attgaatgac	tgaatattt	tttatttgc	agccaaagg	gtgtactga	tatTTTtca	974
gactactaa	aatatatttt	atgaattttt	agtgtgcctt	atcttttttt	gttttttttt	975
tttgatgttga	gtttcactcc	cgttgcctc	gctggagggc	agtggtgca	tctcagctca	975
ctgcaacctt	cgccctcccg	attcaagca	ttcttctcgcc	tcgggtctccc	aagttagctgg	976
gattacaggc	acctggcccc	acacccagct	aattttttgt	atTTTTtagta	gagacagggt	976
ttcaccatgt	tgtcaggct	ggtcttgc	tcctgacctc	aggtgatcca	cccacccctgg	977
cctcccaag	tactgcgatt	gcaggcatga	gccaccatgc	ctggctcgag	gaatattttt	978
ctagttccc	cccaaaaaaa	gcatttttc	tgcaattttt	gttttggttcc	taaagaacagc	978
aagggttaag	gattttaaaa	taatccgtat	tttagatgc	tttctggctt	tgttactttt	979
tatccacagt	agaagtgttcc	agagaatgt	ctccctttt	taatTTTact	ttttggcaca	979
gtatTTTgag	aattataaaat	aatatggaa	tgtttctgg	ctgggtgtgg	tggctcatgc	980
ctgtatattt	ggctacttgg	gaggctgagg	caggagaatc	acttgaacat	gggaggcaga	981
ggttgcagt	agccgaggct	atgccactgc	actccagctc	gggtgcacaga	gcaagactct	981
gtctggaaa	aaaaaaaag	tgtttttttt	ctatTTTcc	accacttgc	gtatgttac	982
taagttactt	ttccctttaa	gtatTTTtg	ctgatgtgc	tgacttaaga	gtatgttac	982
aaaatataat	ttttaaagtt	ctctgaaagc	ccctttaga	gagttttagg	ctatcaaatt	983
gtgtttaatt	cttaacaatt	tttggaaaaaa	ttatagcttc	aatatccgtat	cattccccac	984
aaaaaaagcac	taaaaatcat	gccttgcgg	aggctgcagg	accaagtcat	gttgcattca	984
atgcccattt	tgcacaatcat	gactccctt	caagtagcag	gacagccaca	cttaagaacg	985
agccaaagcca	catggaggcc	gtcatttttgc	gtgacctggg	taagtaacta	tcattttta	985
ttaacttgta	tttagaaggat	ttgagtaaca	tatgtgaaac	ttctgtcata	ggatacagaa	986
ctatataattt	ggaaagtgtct	ttggaaaaaa	tgttattttaa	ataacagctaa	caagtataat	986

87/122

gagggtctca	aacaccagtgc	accctacaac	atttttaaat	gtcgcttcc	tgacagcagt	91560
gatcgttacc	tgcaacgatc	ttattttttt	ttttcatgtt	agtctccaca	cacttgaatg	91620
tagactttt	gaaggcaaaa	tcattgcctt	ttctgagctg	ggagcatgtc	tggcacatac	91680
caagcactca	acagttgatg	tattgacttc	atccagatac	tctggggggc	agtttttcc	91740
tgctactaga	cttccacctt	tcaatgtta	agagcacaaa	tacagagatg	ggcacgttt	91800
ggcattttct	attttgataa	cttttctctt	gtaagatttt	ttaatgttg	aaaaaaaaaa	91860
caaaaaaaga	gggtttaaaaaa	tagtcttagt	tcagatcttg	tgatagaattt	cacacttgc	91920
ttaagctgt	gggcacccctc	ctatcttgga	tgtcatatta	gcttatctac	agcagaattt	91980
ttaactgttt	atgttagtaag	gaagcaattt	tatgtatttt	ttacagacaa	attatctttt	92040
atcttttatt	tttttagacg	gagttctctt	ttgttcttcca	ggctggagat	cagtgtcg	92100
atctcggtct	actgcaacct	ccggctcttg	ggttcaagca	attctctgc	tcagcctccc	92160
aagtagctgg	gcttacagct	gtccggccacc	acaccagct	cattgttttt	tatttttagt	92220
agatgtggg	tttcacccatg	ttggccaggc	tggtcttgag	ctactgacct	cagggtatcc	92280
acccgccttg	gcatccccaa	gtgctggaaat	tacaggcgtg	agccaccgtg	cctggcccag	92340
acaaattatt	atactctgag	tgttagaggc	tttaggatgtt	ttcaacttgc	gctatggag	92400
gaataagtaa	taagatgtg	tacacaacca	aagacccctt	ttcaactatgc	ttcttagatgc	92460
tagtactatg	gatgacacat	ggttataataa	ttgggttagca	tttgcctcta	attactgtg	92520
ctgttacttc	ttcttcaagccc	cttacaggtt	tataattttt	ttcatcaata	atcctctaag	92580
gtatttttt	ttatttgacct	aattttataa	atcaagaaaa	ttaagaccca	gagaagtaag	92640
taacttgtcc	aagatcacat	ggcttataaaag	tggttagagcc	agaatttgc	cccagatgtt	92700
gtgactacat	tgttcttcca	taagcagggtt	caacttctt	gacttggatgc	tgtttcaagg	92760
tcacttctt	agagaaggct	ttgctgacaa	ctaccctctt	gtgccttctt	ccaaggctgt	92820
ccattgttct	agaactttga	atactctatc	tagaataaaag	ctggtctaatt	ttttacagtg	92880
ttttagaaatg	gatctctgac	tgcaaaaagt	ggtctataatt	atctttttat	gttcttagtga	92940
aaggccaaaga	acaagagaag	acctcagatg	tgaagtccat	taaaggtaag	ttctgcctt	93000
ggcagttccac	tgcataaaaaa	agtgtatgtc	tttcgcattt	tgatgttctt	aatccctgtt	93060
tactctctt	tttggcattt	atcattttcg	ccttattttt	taattactt	tgattttgtat	93120
ttatttctt	ctttaaacctg	tataatgttt	taacatctag	catataataaa	gtaggctttt	93180
tttttttttt	tttttttgg	gacggagtt	tgctctgtt	cccaggctgg	agtgcagtgg	93240
cgcgatcttg	gtcaactgtca	agctctgtct	cccgggttca	caccattctc	ctgcctcagc	93300
ctccccagca	gctgggacta	cagggtcagc	gfcgcaccc	tggcttaattt	tttgcatttt	93360
tttagtagaga	cagagtttca	ccatgttagc	cagtatgtc	tcgatctctt	gaccttgc	93420
tcggccccc	tcggccttccc	aaagtgtctt	gattacaagg	gtgagccacc	gcacccggcc	93480
gtaaagttagc	tttttttacc	ttaattttat	tttttgaga	tggagtctt	ctcttataccc	93540
caggctggag	tgcaagtggtg	ccatctcg	tcactgc	atccacccctt	cggttcaag	93600
cgattctctt	gcctcagcc	cccgagtagc	tgggattaca	ggtggccggc	accatggcca	93660
gtaattttt	gtattttttag	tagagacagg	gtttcacctg	gttggccagg	ccagtc	93720
actcttgcac	tcaagtgtac	cactcgctt	ggccctccaa	agtcttggg	ttacaggcgt	93780
gagccaccat	gcctggccat	aagttaggtt	ttactgagcc	ttgtgtgtat	tggctatctt	93840
agtgattaca	gtgaaccagt	gcccttctt	ttaatcacac	atthaatttg	tccctaaaag	93900
tgatttagttc	actttattta	tttagtaaga	caaaaaatga	agaataactt	taactgagca	93960
gtctgtttaa	tgttagaaag	cactgacact	tataaggctt	agttttctgt	catttatcca	94020
gaatgtatgt	tgattacagt	ttttactttt	ttatgttaat	gaacaaccc	aattttaaaat	94080
atattttttt	tattttttgt	ttggatcgat	acattgttctt	tgtttataga	ttagagcatg	94140
ctttttaaag	atgtgtattt	actcactgt	tttatttgc	cagtgtacag	agattgaagt	94200
gggaaaatta	taatggaaat	tgtttccata	gtcattacat	atthaatttc	tcaattttt	94260
tccataaaat	ctgttagattt	ctactttttt	agatttttcc	ttcaatattt	tttatgttg	94320
attgttgcac	ctgttagattt	attctatatc	ctcaatttgc	tggagaagaa	gacttaattt	94380
aacttaggca	agttgtaaaaa	tttagggaaa	aagtaaggta	ccttacagcc	tagtttactt	94440
atttcttagt	taaagccagt	tagattccac	attagttcaa	actgccttct	ttgagaaaaaa	94500
cttgatggc	agtgataaaaag	gcttaaagcc	cttctcaagc	agagacctgt	aaagactaga	94560
tctgactgt	gtagaaggaa	ggaactttaga	tttttcaggc	agtggagaaca	ccagtc	94620
actctaaact	ttggccactaa	cagtatgtc	ttggggagg	gtacttttctt	tcagatctt	94680
cattttgtga	atggggggat	tggcctagat	aatttctaaa	tctctactgg	gctaaaaat	94740
tctgtgttta	tactctgatt	atgaagtaca	taatctgtgc	ttaacattca	ctgacttac	94800
cttaggataa	tacagaagca	gtacaagaaa	cagccccctca	agatgtttgc	agtcttggta	94860
gaaagacaaa	cttatacaca	gaacagttagc	aaatagacca	aaataataat	agctgcccatt	94920
tatagaacac	ttttctcttt	ctgggcatta	gacaaaaactt	gactataacg	gtgaacaaaa	94980
aagacttagg	tcgttgcctc	attgaacttca	cagatttagt	ggggagagaga	acattaatca	95040
agtaatttcca	cagatggctt	agcctagatt	ggtagtgtat	gaagtaaaaag	gatgtgaac	95100

88/122

gacttggaaaa	aaaattcggg	ggcaaatgg	atagaagttt	attattgatt	aaatatgagg	95160
tgtgagagag	aggatattt	aagattgata	cctacccct	ggcttcgccta	acagaaccac	95220
aacaggaaat	tatatgttca	gttttgttat	gttgggtggg	aggtgttttt	gagtccatca	95280
tttatatatg	ttatatatgt	tatTTTtat	gcatagaat	tttaaggct	gagttttaaa	95340
ccaaagggtt	gaggtgatt	tttagagtc	tagcaaacc	aaggtaaat	cctgcctgtt	95400
aaaatggctg	tttactagct	cattaaacct	gggcaagta	tcaacttgt	tttcattttt	95460
gtcttcatct	ctaaaatgag	aaaaatatgg	tcttacaaga	ttgtcctgag	agatagatga	95520
aataatatcc	aaaaaaaaaa	aaggtacata	gagaactcg	tatagtgcct	ggtatatagt	95580
aggtcctcca	ttggtagcta	tcattatca	gttttaacat	agccctcagt	ttgttaatt	95640
agtccaaactg	agtggagcac	tgcaaggaaat	tcagaggaat	ttgagatcaa	caaatgatt	95700
ctgaagttt	ggggaaagct	catggcaat	acacttacct	tgataaaaag	ttgaagaata	95760
agaaagattt	gaatggagag	ttctttctct	tctccctacc	agcccagctt	tttatttgag	95820
gatataattgg	gcaaaaggggc	cttcagacaa	gtagagggag	atTTTACAG	aaagatgtag	95880
atgaaggat	agaaggctgt	aaagaccaga	aaagagaat	gagacagagg	aagcagaag	95940
ccactgttag	tttttgagca	agatattgt	gctgttaagta	tggttttat	gaaaggtag	96000
tctggaaagag	atttgcaggag	tggagaccc	ggaaggTTTT	ttgttataat	acagaaaagac	96060
ttgcactgg	gttggagggt	aaaaataaaa	caggttaagta	aatgttttaaa	catcttgaag	96120
gaaaagtcaa	caaattttgg	caagtaaca	gataacagt	aaaaagaat	ggaccaagat	96180
tttgagttt	ggagactgg	ggattgaaca	gacaggggaa	ttggagggag	aatcagatga	96240
tgatgtttt	agttgtatatt	tagacagat	gtgttggaga	tggttaaagt	aatcgttgg	96300
ggaatgttta	gttagcgat	atcagtata	caagacaaa	gcccaggctca	aagacaagtc	96360
acagatacag	atcagggttt	tttcatctgc	tccacagagg	tgtaaccctag	gagctgttgc	96420
aaacagtcc	tgtgggggt	gtgagtaaga	tgTTTCCCTT	gaatttgcca	gaattacttt	96480
tttgggttt	ttgggttttt	ttctgagaca	gattctcgct	ctgttggccca	ggctgggggg	96540
cagtggcgg	atcgegcagc	tcactgcaac	ctctgcctt	cgggttgcag	tgatttcct	96600
gcctcagcct	ccccaaatgc	tgggatttt	ggcttgc	accaaggccc	gctaatttct	96660
tttgatTTT	tagtagat	ggggTTTcac	catgttggcc	agactggct	cgaactctg	96720
gcctcgtat	ctgcctgcct	cagcctccaa	aagtctggg	attacaggcg	tgaaccactg	96780
cacccgggtc	cttggtaagt	ttatTTTGGT	gggaacaaa	ggaggtttca	gctttttaaa	96840
agtttggaaa	ttattgtct	ggtaataatt	aaagattga	gagtaataat	gtttttagc	96900
agaaagaaata	aaagaagaac	agatagccct	aaaaaggga	gccaaaagag	caggctatat	96960
ctgacacact	gggtgttgc	aatgggtat	aaaaaaatg	agagcaatga	gcagatagaa	97020
gaggaaattt	ggagagtata	ataccatgga	gaccaagaaa	gatagactat	caggaaggag	97080
tggtaaaaat	aagttactag	ttctaaagaga	gatgttaaga	gggaccgggg	aaacccctgt	97140
acaatagtgt	tagtagcatt	ttacattata	taatctaat	taagaaacaa	tgcgagatc	97200
tcaccatcc	tatagactct	tacttgtact	tgtctgaaca	cggaaactgg	ctttgttta	97260
taaataaact	aaaaattttt	ttgtccaaat	ttctctatgg	aaaaaaata	aaaccttctt	97320
taacattgaa	aaaatagttt	gaagacagtc	actttcatt	ttgttaattcc	cacaactatt	97380
attgaatgac	tgaaattatc	tttattctga	agccaaagg	gtgatactga	tatTTTCTCA	97440
gactactaaa	aatatatttt	atgaattttt	agtgtgttt	atcttttttt	gttttttttt	97500
ttgagatgga	gtttcactcc	cgtgtctag	gctgggggg	agtgggtca	tctcagctca	97560
ctgcaacctt	cgcctccccc	attcaagca	ttctctgc	tcggctcc	aagtagctgg	97620
gattacggc	acccggggcc	acacccagct	aatttttgt	atTTTtagta	gagacagggt	97680
ttcaccatgt	ttgtcaggct	ggcttgcac	tcctgacctc	aggtgatcca	cccaccttgg	97740
cctcccaaag	tactcgatt	gcaggcatga	gccaccatgc	ctggcctgag	gaatattttt	97800
ctagggtccc	ccccccccca	gcatttttgc	tgcaatttt	gtttgttcc	taaagaacag	97860
aagggttaag	gattaaaaaa	taatccgtat	tttagaatgc	ttctctggct	ttgtactttt	97920
tatccacagt	agaagttctc	agagaatgt	ctccctttt	taatTTAact	ttttggcaca	97980
gtatTTTGT	aattataat	aatattagaa	ttgtttctgg	ctgggtgtgg	tggctcatgc	98040
ctgtatccct	ggctacttgg	gaggctgagg	caggagaatc	acttgaacat	gggaggcaga	98100
ggttgcagt	agccgaggtc	atgcactgc	actccaggct	gggtgcacaga	gcaagactct	98160
gttggggaaa	aaaaaaaaaa	aaaaaaagag	ttttttttt	ctatTTTCC	accactgtat	98220
taagttactt	ttccctttaa	gtatTTTGT	ctgtatgc	tgactttaaga	gtatgttac	98280
aaaatttaat	ttttaaagg	ctctgaaac	ccctttatga	gagttttagg	ctatcaaatt	98340
ctgtatccct	tttaacaatt	ttttgaaaaa	ttatagctt	aatatccgt	cattccccac	98400
aaaaaaagcac	taaaaaatcat	gccttgc	aggctgcagg	accaagtcat	gttgcataat	98460
atgccttc	tgcccaat	gactccctt	caatgcag	gacacccaca	cttaaaagac	98520
agccaaagcc	catggaggcc	gctcatttt	gtgacctgg	taagtaacta	tcattttta	98580
ttaacttgc	tttagaaggat	tttgagtaaa	tatgtgaaac	ttctgtcata	ggatacagaa	98640
ctatataatt	ggaaaagtgc	ttggaaaaaa	tgttattttaa	ataacagcta	caagtataat	98700

ggtagctgt	gttgtgttcc	tgttaaatata	gaatataaag	catgcccagt	agaaaaaaacaa	98760
gcattccag	aagaatata	tctgatca	aatataaaat	atatgaaaaa	gatgtctcac	98820
tttattactg	agggaaatgc	aaattaaaat	aatcagttaa	tgttctccta	acacattagc	98880
atattttta	aagtttgaca	atttgaatgt	cagtgaagat	gcagggaaat	acccttccta	98940
tttagtgata	atataatctg	gtgaagactc	tttggaaagc	aatttggaaa	tcaagtataaa	99000
atatgcatgt	catttaggcc	actctttctt	agacactagcc	ctcagatatg	ctcattcata	99060
tgtcagggt	tgtatgtgt	tgtgtgtgt	tgtgtgtgt	tgtatatgt	tgtatgtatg	99120
tatgtatgt	tgtatgtgt	aggctattca	ttatagattt	gttggata	gcaaaaaatt	99180
atggacaaca	tataaaatatc	tgttataggg	aaataaccaa	attgtgttat	acgcacatgtc	99240
tggagtataa	tatagccatt	tgtttctatt	tatttatttt	cttggagacag	ggttttactc	99300
tgtgccccag	gctggagggtc	agtgttatga	tcatggttca	ctgcagccct	cacccctctgg	99360
gcacaagcca	ttctctcgcc	tcaagttcca	gaggtaactag	gactgcaggc	atgtgtcacc	99420
acacccagat	aattttttaa	tttttgttag	agacagggtc	tcactatgtt	gcttaagctg	99480
gtctcaaact	cctggcctca	agcaattctc	ccacacaggc	ctcccaaagt	gtctgggatta	99540
ccaacgtgaa	ccaccacacc	tggttcagtg	tagccattta	gaaatctaaa	aaagacgtgg	99600
gaaaatgtct	aaggcatgtt	taaattgtgag	aaaagcaagt	cacagtatgc	atggtaaaat	99660
ccgttatatt	aaaataatgtt	cttccaaac	aaaacatata	cgaggagacc	tttattttgt	99720
cagattttct	tacccaaat	tctgcactta	gaaaattgtca	tgtcatgtt	tcataagttg	99780
aaaaaaagat	ccatgaacca	atggacttct	aataaaatca	gtcctgtttt	tgacatctct	99840
ctctactttt	gtgtatattc	aaaccagagt	gtcaatgtgt	tttgtggca	cacttagcaa	99900
taatacatag	cagacaaaat	gcatatagct	cagagagtaa	aattgttaat	tttgcttagat	99960
cactcataaa	ttgctgtatga	gaattttaaa	tggtgccatgt	gctctggaaa	acaggcaggtt	100020
tctttttttc	tttttttttt	tcttttttgg	acagggtctc	actctgttgc	gcaggctgg	100080
gtacagtggc	gtgattacaa	ctcaactgcag	cctcaccctc	ctcagggtca	gttgatcctc	100140
cctcagtctc	ctgagtagct	gggactata	gcatgcacca	ccacgcctgg	ctaatttttg	100200
tatttttttt	tttttttttt	gttagagacgg	gttttcgcca	tggttcccag	gtctggctca	100260
aactccgtga	atcaagcgat	ccacttgcgt	aggcctccca	aagtgtgg	attacggggcg	100320
tgagctactg	tgcctggc	aggcagttt	tttgggttgg	tttgggttgg	tttattttatt	100380
tgttagacgg	gtctcacagg	ctggagtgc	gtggcccaat	ttttggctca	ctgcaacactc	100440
cgcctcccg	gttcaagcta	ttctctgtcc	tcagcctcc	gagtagctgg	gatgacaggt	100500
gcctgcccata	atgcctggct	gattttgt	tatttttag	atatggggtt	tcaccatgtt	100560
ggtcaggctg	gttttgaact	cctgaccctc	ggtgatcgc	ccgcctccgc	ctcccaaagt	100620
gctgggatta	caggcatgag	cgtcataccc	ttggctgtgg	tttctttaga	cgtgaaacat	100680
gcaattacca	tatgaccttg	cagttgcact	ctgttattat	cccagataaa	tggaaaactta	100740
ccttccaata	aaaacctgtg	cacaaatgtt	catagcagct	taatattgaa	aaactggatg	100800
tttcttcagca	gttgaatgaa	ctgggttatt	cataccatgg	aataccattc	agcaataaaaa	100860
aggaacaaac	tgttgatata	ttttaaccacc	tggatgtaata	tcaaggat	tatgctgtca	100920
gacaaaaacc	agtccctata	gactacatata	agtatgattc	cgtttggata	atatttttga	100980
aatagagaaaa	ttaagagaaaa	tgaaaagatt	agtgtttgc	agatgtttaga	gacaggagg	101040
tgagaggggt	aagtgggtgt	agttataaaa	gtgcaacatg	agggatctt	gtgtgttga	101100
agttgtatct	tggcagtgg	tgcagaaatc	tcaatgtat	aaaattacaa	agaactaaaa	101160
acaagaatga	gtatagataa	aactggggaa	atctgaaacaa	gttagagtgt	tgtatcaactg	101220
tcagttatct	agagtgtat	tgtactatag	tttgcata	tgttaccatg	ggagaaaacta	101280
aagtgtacaa	gggatctctt	gttattttt	ttttttttgg	gatgggtttt	cactatgttc	101340
cccaggccgg	tcttgaactc	ctgggctcta	gtgatccgccc	tgccccagcc	tcctaaagta	101400
ctggattac	aggcgtgagc	gaccatgcct	ggccctttca	gtattgtatc	ttagaacttc	101460
atgtatct	agcattatct	catagaat	attttaaaaga	aattgtaaac	ctccacagaa	101520
atcagaattt	cctcaagttt	gtgtatgtt	caaagatgaa	ctagttgaca	tcgacagttaa	101580
gactgaggat	gaagacacca	cgtgttca	aaaatgtt	tgtatata	tggattaa	101640
agaactctt	tgacaaattt	atgaaaccct	cagtca	tataagaatg	ccatcttta	101700
tgatcatgt	atgaaagcca	atttttaaa	aaattttttt	tcttttctta	caatttagctt	101760
gtggttataa	tttttttttt	gttaatata	agataaaatgt	tttttttttt	agtttagttt	101820
catttttcaa	ggtacgtat	caaagactat	ttttaaccct	ctatgtatgt	ataatgtctga	101880
gttcataaca	tctttgtata	tatatccaca	attttccctc	aggataatgt	cctacaagt	101940
gaatttactgg	actgaaaata	atgcagttt	ctaagacttt	gttatctgtt	cctgaatgt	102000
cctccaaaaa	ggttttggca	gtttacatcc	tcatgaccag	cgaatgagag	tgttgcctat	102060
tttccctgtgc	ccttggtaact	gtttaataat	ttttggaaaa	aatctaaat	gacagacaaa	102120
aatgcatttt	atgtttaat	gttttctgg	gattttttat	gagggtttagt	atagttttta	102180
atatttttat	tggcccttt	ggaacttaga	tcataagttt	ttttttctta	gaatttatgt	102240
agtctgggct	gggcgcagtg	gctcacgcct	gcaatcccag	cactttggga	ggccgaggtg	102300

90/122

ggtgtgattgc	cgaagggtcag	gagtttgaga	ccatcctgac	caacatggtg	aaaccgaatc	102360
tctactaaaa	gtacaaaaaac	tagtcagcg	tggtggcggg	tgcctgtaat	cccagctact	102420
tagaggctg	agtcaagaga	atcgcttga	cccgggaggt	ggaggttgg	tgcattgagc	102480
cgagatcgcg	ccattgcct	ccagcctagg	caacaagagt	gaaaagtctc	aaaaaaaaaa	102540
aaaaaaaaaa	aaaaaaaagaat	ttacatggtc	tgaattgcca	ttaaaagaga	tatgagaatt	102600
attgagtaac	aaataactt	ttaataattt	aggcaagttt	tggacgattt	tactttgttt	102660
agaaaccaaa	agcatagtt	ttgttagttt	tttattttt	tttagtgc	ggaagtaaaac	102720
tttatttcaag	gtctctggta	ccagttgtt	ccaaaagtga	ttgactaattc	tgtcaatctg	102780
aaatttattt	ttgtgcact	gtctaatttctt	ttgcttctat	cttttaggca	gatcttgc	102840
ggactaccag	actcaagaga	ccaaatcaag	ccttcttaag	acccttgcac	aagtcttgc	102900
cgacactatt	gtcctccctt	acttcattca	attcatggaa	cttcggcgaa	tggagcattt	102960
ggtgaaattt	tggtagagg	ctgaaagttt	tcattcaaca	acttgcgc	gataaagagc	103020
acacagtcta	aacacagtga	agcagagctc	actggctgag	cctgtctc	catctaaaaaa	103080
gcataaaact	acagctgtt	ttttactga	ttctcttgat	aagagatttg	aggattctgg	103140
ctcagcacag	ttgttatga	ctcattcaga	aggaatttgc	ctgataataa	gaactaacag	103200
cactcagaat	cacttgcgc	tttccagga	atgtgacagt	gcccatctt	tccgtcttga	103260
aatggccaga	gcaggaactc	accaagttt	catggaaacc	caagaatctt	cctctacact	103320
tacagtagcc	agtagaaata	gtccccgttc	tccactaaaa	gaattgtc	gaaaactaat	103380
gaaaagttag	tatgtgat	tctttgtgt	acatatgtt	ctcactttt	tttttaattt	103440
tactaaggcag	aacttcagat	gaggaataaa	atgattggaa	tatttttttt	ctcctctaa	103500
tacttgtaaa	tttggagaa	tttggagat	gttagtagat	cagatcagt	tatggaaaag	103560
gagcaggagt	gactggacct	tctaagaagt	gtgttatcag	aatttagaaa	tgaagggtc	103620
aatgtcctac	ttttcccttc	cactgatttt	gacatcaa	cattatccac	atagccttat	103680
ttcctccctc	ggtcttaattt	tttataatat	tttactgcac	tttgcagata	aaatttttaa	103740
aaaattttta	aaaattggca	ataagtgaca	tttattaatgt	tcagtgc	gtgttatattt	103800
ggattttattt	tattagtcac	aagacctttt	tgccaggat	aggcatgatt	atcttttttt	103860
ttttgagatg	gagtcttgct	ctgtcgccca	ggctggatgt	caatggcg	gtctcggttc	103920
actgcacact	ccgggttcat	gccattcttcc	tgccctcagcc	tcccaatag	ctgggactac	103980
aggcgcctgc	caccacaccc	ggcttatttt	tttgcatttt	tagtagagac	gggggttcc	104040
catgttcgccc	aggatggct	cgatctcttgc	actttgttat	ccgcctgc	ccgcctccca	104100
aagtgttggg	attacaggca	tgagccaccg	cgcccgaa	gattatctt	tttacacatg	104160
agaaaaaccag	ggcttagaaa	ggtttaggtaa	cttccctctag	gttgcacat	aatgtggac	104220
ctagaagcat	tttgacaa	gcacacttgc	tttttcttc	tctttagttt	tagaaattat	104280
atactttaa	tttacacttgc	ggatttgtat	tagacacttgc	tcatgttctt	tttcatcttta	104340
aatgttctt	gtgtctttaaa	gggtttagtgc	atttcttgc	atcttttagt	tcactcattt	104400
tcagtgaact	aaaatgaggat	ctaatctgc	actgaatcaa	gttttcagca	tggttatttcc	104460
ttcctccctc	cctccctctt	tccttccctc	aaccaggctc	ccggaggat	gggattacag	104520
gcgcgcgc	ccactcttgc	ctaattttta	tattttagt	gagacgggtt	ttcaccatgt	104580
tgttcaggct	gatcttgcac	tcctgacctc	aaatgttgc	cctgcctcg	cctcccaaag	104640
tgtctggatt	acaggcatga	atcaccacac	ctgacggcat	gttattttca	tcgcaagtt	104700
actgttaagct	gggagaagtg	gcacacactt	gtactcccag	ctactcagga	agcttaaggt	104760
gagaagattg	cttgagccca	ggagttttga	gaccaacctg	ggcaacacag	caagacccca	104820
gctcaaacaa	agaaaaaaag	ttatttgcatt	tttttatttt	atggatcatt	ttttgttagtt	104880
tcttattctt	ttcacccatc	atccccatc	tttgcattat	tttttttttta	tttagttttt	104940
ttaaaatgtat	atttgcgt	taatttgc	atctacatgt	ttttgtggac	ctgactcagc	105000
atttcttgc	tttgcgt	tcagactgtt	ggtggcttgc	gatttttagt	atttttggcc	105060
gtgaacatgt	ttttggact	tttgcgt	ggaatttctt	gtgtactctg	tataaattaa	105120
gttacttcag	gtgttttgc	tttttttttgc	ccatgcac	ggggcctgg	tactaccct	105180
tctgttacca	ttttaaaact	aattttgc	tttgcattat	gtactgatcc	tgtatgtat	105240
caggtttata	cttgcgt	aaatatggt	tttgattat	gggttattgtc	ccagatgggt	105300
ctggagtatt	aatatgc	ctttttaact	taatgttttgc	tccctgtaaa	actccaaat	105360
tctgaattcc	agaatactac	tggcccaaaa	tgttttaagat	aaggcacttgc	cctgtatttgc	105420
tttctgcctc	ccactat	ccttagttt	acacaaactc	acctttttaa	aaaacatttt	105480
gagagaattt	agatgggtt	agatgttca	accttgcattt	ggaaatggaa	gtccaaagtc	105540
tgtttctgt	attttttttt	tttgcgt	gagtctca	ctgtcacc	ggctggatgt	105600
caatgacgt	ctctcagctc	actgcaac	ccacccccc	ggttcaagcg	attcttgc	105660
ctcagcccc	ttagtgcgt	ggattacagg	tgcacc	catgcctggc	tgatgtttgt	105720
atttttagaa	gagatgggtt	ttcgccatgt	ttggccaggct	ggtcttgc	ac	105780
gtgatctgc	cacccatc	tccaaatgt	ctaggattat	ttttgtgtaa	ttgtataaca	105840
tttatttgc	tttagaaact	tcttgc	agtggtaatt	ttcaataaaaa	atagaaatag	105900

91/122

cagtggagtt	attaaaagag	cattagttac	attttccct	ttttcattat	cttcaaataat	105960
tatataatgt	aagtttgacc	ttttaaaat	gtataacttgt	atcagttta	acacatacat	106020
agattcctgt	aactgtcacc	actataaggg	taaagaacag	ttagttcctt	cacctttgaa	106080
gtcaagcccc	acctctatcc	caacacttgg	caaccgctga	tctttctccg	tctcaatagc	106140
tttgccttt	ctctttttt	ttcttatttt	ttttttgag	acagcgtctt	gctctgtcgc	106200
ccgagctgga	gtgcagttag	gcaatctcg	ctcaactgaa	cctccgcctc	ctgggttcaaa	106260
gcagttctcc	tgcccttagcc	tccttagtag	ctgggattat	aggcacgcac	caccacaccc	106320
ggctgatttt	tttgatattt	tagtagaaat	gggggttccac	catgtggcc	aggctggct	106380
caaactcttg	acctcaagtg	atccacctgc	ctcggcctcc	caaagtgtcg	ggattacagg	106440
cgtgagccac	tgtgccaat	caggactttt	ttttttaaa	tttacattca	acttgtcatt	106500
tttttcttgt	atggattgt	ccttcagagt	cacacctaag	agccctttgc	ctaagcaaag	106560
gtcatgaaga	ttttctcata	tgttccctt	taaaagtatt	gtgggtggcc	aggtgccatg	106620
gcttatgcct	gtaatctcg	caacttggaa	agctgagggt	ggcagattac	gaggtcaggaa	106680
gatcgagacc	atccctggta	atgcggtaaa	accccatctc	tactaaaaat	acaaaaaaaaa	106740
aaaaaaatta	gcccggcgtg	gtggcgggca	cctgttagtcc	cagctacttg	agaggtttag	106800
gcaggagaat	agtgtgaacc	cgggaggtgg	agcttgcagt	gagccgagat	cgccgcactg	106860
cactccagcc	tgggcaacac	agtgagactc	catctcaaaa	aaaaaaaaaa	agtattatgg	106920
tttacactt	tacggtttaga	tatatactt	ttttgagttt	atgtctata	atgtgggg	106980
ttacgtcaga	ttttttgtt	tttggattt	tttacatatg	gatgtctagt	tgttctataata	107040
ccattttgtt	aaaagacaac	ctttaactcca	ttgaattgcc	tttgtacttt	tgccatattt	107100
gtctaggcct	gtttttggac	tcctttttct	gtttcatgt	gtgtgtgtct	attccctttgt	107160
taataccaca	tggtcttaat	tactgtatag	taagtcttaa	aattgggtaa	tgctggcctt	107220
ataaaaacgaa	ttggggaaat	tttttttttta	ctcttatttc	cttttcttag	aagagattgt	107280
gtagaatttgg	tgtcatttct	tctttagata	tttgggttggaa	ttggggaaatg	atgcccattctg	107340
ggccttaggt	ttttttttt	gtgtgtgaga	cagagtctca	cttctgtcac	ccaggttggaa	107400
gtgcagttgg	gagatcttgg	cttactgca	cctctgcctc	ccagggttcaa	gttattcctcc	107460
tgccctcagcc	tcccaaataag	ctgggattac	aagcgtgtgc	cacatgcgcc	gactaattttt	107520
tgttatttttta	atgcagacag	ggtttccacca	tgttagccaa	gtctgtctcg	aacttgtgac	107580
ctcaagtgt	tagccccat	tggccctccca	aagtgtttag	attatactatg	tgagccacccg	107640
tgcttggcag	gggccttaggg	ttttttttt	cagagtattt	taaactatga	attcagattt	107700
ttaatagat	ataggactat	ttaagtttac	tgtttcttct	tgagtgaattt	tttactgttag	107760
tttatggcct	ttgagtaatt	aatttttattt	aattgtcaaa	ttttagtgcg	tgtaattattt	107820
tatagcattt	cgggtttgt	gtggatctcc	tcttttttttct	ctgggttgg	caattgtgtc	107880
ttttttttct	ttgtcagatt	gtataggat	ttttagtct	tttcaagaaaa	ctagtttttgc	107940
ttttgatttt	tctgttgtt	tgtttcaat	tttatttattt	ttctgtcttt	tattattttct	108000
tttctattat	ttctgtcttc	tttgggttta	ttttacttct	ttttttttct	ccaagttgt	108060
taaagtagaa	acttagattt	ctgggttttag	acctttcttt	tctaaagataa	gcattttaaata	108120
ctgtaaattt	ccttcttaacc	actgttttag	ttacacccccc	acaatattctg	gtattttggaa	108180
ctgagcacaa	atgaaatgtt	cttattttcc	ttgaatctta	tttttttacc	aatgaatttat	108240
ttagaaatat	tttattttatgt	tttcaagcaaa	ttggagactt	ttttccctgtt	atttttcttac	108300
cattatttc	tcattttcatt	atattatggt	cagagaatat	attttgaatg	atttcatttt	108360
ttaattttta	aaaataacat	taaaaaatgt	ttttaaaatgt	gaatataccca	catacagttat	108420
aaagattgtt	cattctttt	tttggacagttt	ttcttataat	gtcaatttgc	tttagttgg	108480
taatgtatgt	gttcagttt	tctttttttct	tgctgataact	ttgtatgcag	ttatatactact	108540
tttattttact	gaagagtttt	gaactttccca	actacaattt	ttttttccaa	tttttttttcc	108600
agctctatct	ggttttgttt	catgtttttt	gaggctctgt	tgtaggtgt	gtacacattt	108660
aggatgatat	cttctgggtt	aatttgcctgt	tttattcatat	tgtaattccc	tcttttatggt	108720
aattttctt	tttcttaagat	cagaaaatatac	tgtttttccaa	tttataataga	cactgcagct	108780
tttattttat	tagtgcgtt	atggcatatata	ttttttttttt	ttttttttttt	tgatctaccct	108840
ttataattct	atttaaagggg	ggcttcttgc	aggcagcata	tagttggta	gtgttatttt	108900
tttattttat	tattttat	tttattttat	tatttgcata	gagttttgtt	tttgcgttgc	108960
aagctggagt	gcagttggtc	aatccctggct	taccacaaacc	tccacccctt	gggttgcagt	109020
gattctctgt	cctcagccct	ccaaatgtat	gggattacag	gcacgcgcac	catgccttgc	109080
tgatttttttt	tattttttatgt	agaaaacggat	tttcaccatgt	tttagccaggc	tcgttatttt	109140
ctcctgaccc	cagggtatcc	acctgttttgc	gcctcccaaa	gtgttggat	tacaggcgtt	109200
agccactgca	cccggtctgg	tcatgtttt	tttaatcttt	tctcacaata	cagggtttttt	109260
gttggtaat	ttaattttat	taatataaaat	tttagtataa	tttatttacat	taaatgttaac	109320
tgttgcactg	gggttattttat	aatgtgtaaa	tataattttt	ggttataata	taattttatatt	109380
actcataata	atattttat	ctttttttttttt	agattttttttttt	tttagtataat	ttttttttttttt	109440
tttccctct	ttgattttttttt	ttttttttttttt	ttttttttttttt	ttttttttttttt	ttttttttttttt	109500

92/122

tagtatttgt	tgatcattct	tgggtgtttc	ttggagaggg	ggatttggca	gggtcatagg	109560
acaatagtgtt	agggaaaggtc	agcagataaa	catgtgaaca	aggctctgg	tttcctaga	109620
cagaggacc	tgcggccctc	tgcagtgtt	gtgtccctgg	gtacttgaga	tttagggagt	109680
gtatgactc	ttaacgagca	tgctgccttc	aagcatctgt	ttaacaaaagc	acatcttgc	109740
ccacccttaa	tccatTTaa	cctgagtgg	aatagcacat	gtttcagaga	gcaggggggt	109800
gggggtaagg	ttatagatta	acagcatccc	aaggcagaag	aattttctt	atgacagaac	109860
aaaatggagt	ctcccatgtc	tacttcttt	tcacacagaca	cagaacaaat	ctgatctctc	109920
tttctttcc	ccacattttc	ccctttctt	ttcgacaaaa	ctgccatctg	catcatggcc	109980
cgtttcaat	gagctgtttg	gtacaccc	cagacgggg	ggcagctgg	cagaggggct	110040
cctcaactcc	cagatggggc	agccgggcag	aggcgcccc	caccccccag	acggggcagt	110100
ggccgggccc	aggcgcccc	cacccccc	ccggatgggg	cggtggcc	ggcggggggct	110160
gaccggggcc	ctcccccc	gacggggcgg	ctgggggg	gggggtgac	ccccccaccc	110220
cctcccaat	ggggcggctg	gccccgggg	ggctgggg	cacccccc	ccggacgggg	110280
cggctggcc	gctgagggg	tcctca	gcagacccgg	cggtggcc	gcccgggg	110340
tcctcaactc	tcagacggg	cgccgggca	gagacgctcc	tcacccca	gttgggggtgg	110400
cggctggcc	gagacactcc	tcagttcc	gacggggtgc	cgccgggca	ggggcgtcc	110460
tccatccca	gacggggccgg	cgggccagag	gtgggtcccc	catctcagac	gtggggctgc	110520
cggcagaga	cactccteac	tccttagacg	ggatggcagc	cggaagagg	tgtctctcac	110580
ttcccaagac	ggggccgg	tcagagggc	tcctcacatc	ccagacgatg	ggggcttagg	110640
cagagacgt	cctca	cgacgggg	ggcggccgg	cagaggtgc	aatctcgcc	110700
ctttgggagg	ccaaggcagg	cggtgggaa	gtggaggtt	tagggaggt	agatcacggc	110760
actgcactcc	agcctgggc	acattgagca	tttagtgc	gagactcgt	ctgcaatcct	110820
ggcacctcg	gaggccgagg	caggagatc	actcgccgt	aggagtcg	gaccagcccg	110880
gccaacacag	cgaaaccc	tctcacc	aaaatgaaa	aaccagtca	gttggggc	110940
gtgcgcctgc	aatcccaggc	actctgcagg	ctgaggcagg	agaatcaggc	aggaggtt	111000
cagttagccg	agatggccg	agtacagtcc	agcctcggt	tccacaactt	tgtggcattc	111060
agagggagac	cggggagagg	gaggggaga	cgagggagag	ccccctttt	gtttttttt	111120
ggatttttgc	aatttttct	taattttt	tatcttact	attattttt	ttttttgagt	111180
gattctctg	ccacagctc	caagtagct	ggactcgagg	catgtccac	tacacccc	111240
taatttttt	gtattttt	tagagacagg	gtttcacat	attggccagg	ctggcttga	111300
acttttgacc	tcaagtgtc	cacctgcctc	ggcctcccaa	agtgtgg	ttacaggcgt	111360
gagccaccat	gccctgcctt	ttcttagaa	ttatatttt	agttttgt	tgtatctttt	111420
tattaggct	tttttagtgc	ttcttagga	attacaat	aatactttt	cacagtgtac	111480
tcacatccaa	tatTTgtaa	cttcaagtgg	atgtagaaa	acttaaccac	cataaaaata	111540
gaactaggga	tgaggttaaa	aaagagagag	aaaagaaaatg	taataaagat	ttaataaacac	111600
cgtttttttt	tttttttctc	ttttttttt	gagacagagt	ctctcttct	gttaccaggc	111660
tggagtgcag	tggcgtgatc	ttggcgtact	gcaacctccg	cctcctgg	tcagatgttt	111720
ctctgcctc	agcctactga	gtagctgg	ttacaggtc	gcccacat	gcccagctaa	111780
ttttgtttag	tttagtagag	agggggc	tgtgtggcc	aggatgtct	cgatttctt	111840
accttgcgt	tcgtctct	cagccccc	aagtgtgg	attacaggcg	tgagccacc	111900
cgcggccta	agtctttaaa	tatTTttt	acattgcact	tttctctt	tcctcttctt	111960
attttagtaa	cccaaatgtt	agtttggta	ttgtttggca	gttctctt	gtttttctt	112020
cttctttaaa	tttttttctc	ctgtgttca	gtttcgaaaa	tttcttattca	tctgtcttca	112080
aattca	ttttttcc	ttatTTccat	tctgttattt	agtcttgc	gttatttttta	112140
aattttgtt	attatgttt	ttagttctaa	aattttctt	ttttgtgtat	gtcttataact	112200
ttgtctctg	aactcttatt	tgtttcagga	gtatctt	ttcttagagc	atggttttag	112260
tagctactt	aaattttgtt	tatcatccc	gcatatgtt	cctcttattt	gtctttctc	112320
tttgagat	atgggat	ctgtttttt	atatgacaat	taattttgg	ttgtatctt	112380
gacagtgt	tttacgttac	atgttctga	atcttgc	aatctgtgg	aaaatattga	112440
agttttgt	ttaacaagca	gttgcac	ttagttcag	tccacaaatt	ctaaggcagca	112500
ttctgtggc	tctgggttca	tcatcgtt	agttttgtat	tttatctgt	tatgtgcctt	112560
tctgtgttca	gtctgggacc	tggccaatgg	tcaggcc	aaggccctt	acacttttag	112620
aacaggggcc	atgcacaccc	agctcacag	tggcccccgg	agtgcacata	caactcgac	112680
tttcatggg	ctccctt	tctgtatgt	ccctgacac	ttctgcctt	taagaaccc	112740
ccttatccc	tttctgtt	tctgttctg	aaagtca	tttagattcc	ctataacttca	112800
gcacactcc	tgttagctat	tcaac	ttggccac	ttcttctt	ttggactgc	112860
gtttcttctt	tcagaaagta	ggatttttgg	agctgtgtc	attgtgt	ttggctgtct	112920
gatgtgtgc	gggagtgc	ggagagaaag	gaacaaaaca	aaacaaccc	ggggatttcc	112980
tccactcttct	ttgatccgt	agagccccc	ttctgttcc	tcagacc	aatagaggc	113040
ctgtcttgg	acttcttctt	tgtgc	atgtgc	tcagttttt	gttccaggcc	113100

aggaggtgct	ggacaaactt	gtcaggagta	cggaggtact	gcaaggttctg	attacttttc	113160
tcagtccacc	tgcttccaag	tccttggatg	catttgcaca	ttgtttttag	ttgcattcca	113220
tggagagac	agaagagtgt	gcttatttca	tcttgacata	cttatttagga	tttcatatca	113280
aatcaacgga	tgatattctc	tatattaatt	tgctgtttc	ccttttagcaa	gcacattagg	113340
aaaataaacac	ttaacaccc	gccttggtg	gtttctgtca	taattattaa	tacttgactt	113400
ttttttttt	tttgagacgg	agtctactc	tgctctttga	ggcattgtcc	ccataaaactt	113460
ttggtaaaagc	atcaataatt	ttatcttca	tccacacaag	cttcaccata	aattttagtgt	113520
ttattcttc	attttagcag	aattcatgtt	gctccaatag	gggctgtctt	caaactgtatg	113580
ttttccctt	cttagtgctc	cagagtagat	cctgttcaga	tacgttataa	caggttaata	113640
tgagtttatt	ttggtgtaaa	agtactttga	aattcatgca	tagtttttc	atcatatgca	113700
ttttccatag	ctttgaacac	ccccatgtaa	ctctcccttct	ccacaaacca	aacaatgaaa	113760
aagcacctt	gtgatggaaag	tttattttgc	aattttagact	cacagtgtatc	taagccctgc	113820
tattcatgaa	tataattcat	tactggatc	caagttgtt	tttgggtttt	gaagttctct	113880
tcttccctt	caggatatag	acaagatgca	gtgaataactt	ttacaaaata	tatatctcca	113940
gatgctgcta	aaccaatacc	aattacagaa	gcaatgagaa	atgacatcat	aggttaagcag	114000
tgcttgaacac	tatggcaaaa	aaaaaaatgac	aaaaaaatgca	cagaactgac	aattttcggtt	114060
attgactaaag	ataatttttt	cttaacatgg	aattttagac	ttcccttcct	aatttggttt	114120
ctgagtattt	tttatatctgg	attatagctc	attttttaaag	tttctcggt	gcattcggtg	114180
cgagggtctt	tgcttggggc	agatggctg	cagtgttagc	ggtgcctagg	cctgcccgc	114240
gctgagcagc	cgggccggcg	ggcgctacg	ctaaccggca	cagaccaccc	gatggactgg	114300
ccggcagccc	cgcaccagtg	cacgaagtgg	gcgggacaga	aacttctggg	gttggaaagtc	114360
cagtgaggct	aaaagccggt	acccaaatgt	ctaggcatca	gggctgcagc	ccaaagagtct	114420
cacgaccagt	ggcacaactgg	atggccagac	agggtgtctca	gtggggcct	ctccgtctca	114480
ggcttccatc	ccacttctca	gtgggcttga	cgtccctggg	caccctggat	gtctacactgc	114540
attagccaga	gccccatcacat	ggcctgtgac	ttgccttttt	ttgccagtt	attgtgccac	114600
acacagtgtc	atttctgtgt	catttggcac	agctggaggt	gcaaggagga	gggcagccct	114660
atgtccagtc	ccagtttccac	gtaaactttat	tcttctgtaa	aaagacaaat	tgttaacacctt	114720
aaaaaaaaaa	aaaaaaaaaa	agtttttttt	atatgttgg	cccaattttt	taggttttaa	114780
cctgaaataac	aatgacagac	agatcaataa	atagacaca	tttattaaac	acttactgtg	114840
tcccagacaa	tattccaaacg	actttttatg	gatagactca	tttaacttc	taaagaactt	114900
tgtggataaa	atacagttat	ttttagatg	aagaaaactga	agcacagaga	agtttaagtgc	114960
tttgcctcagg	gtaacagctc	agatatggca	gagtcaaggat	ttgaaactag	accctcacat	115020
accttaactg	ctgtgctgt	gcagttttt	tcatactgt	gttggggacc	acccttctct	115080
tatgccctca	ccccctggca	aaaaaaaaaa	aaaaaaaaaa	aaatatatat	atatatatat	115140
atatatatat	atatatatata	aatataaaaat	atataaaaat	atatatatat	ataaaatata	115200
tgtatttagta	tatatgcata	tatagtatata	attatataatt	agtatataata	ctaatatata	115260
atatacatat	tagtgtgtgt	atatatataat	atactagaat	aaaaaaaaatca	aagtatctca	115320
gagtagtaag	gacaaacatt	tcagaaaaat	gttttcattt	tatatacatg	tagttagtgc	115380
tatgtgttt	caacaaatatt	ttttttttata	ggttatacga	aaagattttgc	aaagctttta	115440
ctgtgtttt	tcagggagac	cttaggtgaa	cgtatattca	cagataaaag	agtttattta	115500
ttcattcaat	aaatattaca	ttctcataag	tcctaatatt	atgtttttt	atttttcaaa	115560
aaagtttagta	tttgcatttt	atgaaataag	acatgttctt	gcacttttag	cagatctgtc	115620
ccgatgtgg	gctttttttaa	tcctttagtgc	gggtgttttg	cactcaatca	ctgctggggaa	115680
cagaagacc	cctgttagtgc	tcaggtgttgc	tttttttttt	ggccctactgt	accctccatg	115740
tagtatttct	ggggtccatg	tcatgttggc	tccattttttcc	tttttttttt	cccacacacaga	115800
tacctataac	ggctataaca	taggcctgg	ggctgttgg	ggcttattccc	tatctgttttgc	115860
tattaaggg	gtactgtttc	actgagttt	gctgacagat	gttgcattga	gatttgaggt	115920
tttctgtttt	gttgcattat	ttttatgtgg	gaattttgtca	ctatcatcat	cccttagaccca	115980
gcttttctta	gtatataac	agggtatgttgc	tgactgttca	gagtttgcct	tttggaaagaa	116040
ttgttggct	atgtttttttt	ttttgggggg	agttctgtacc	agtttaatgc	ctgactggcg	116100
tgtggataaa	aaggaagcag	tttcaagtca	aataaaacac	ttaaaaatgaa	accacactgc	116160
aactctttt	cttttactta	agcttaatca	aattaatgtat	gatgtatcc	catgaaggaa	116220
aagtcttcgt	aaggatcaag	ttgataacat	tttgcattca	aagaatttgc	gaaaacacctt	116280
atccctgtt	ctatcattat	atatttttag	atgttaattt	cctgtgttgc	tttggcaag	116340
tcatttttcc	tccttgcggcc	ccatctttaa	tcctgtccaa	attatttgc	tccttcttgc	116400
gttggactat	ttaatatacg	ctgttccatca	agttagtttt	gttcaaggaa	gccttcactt	116460
tagtcttac	tgtgttacca	ctttgcattag	tcttgcatttt	aatgtatcc	ttggattttt	116520
gggttgcata	actaattact	gtttttatgt	gaggattttag	agtgtatcc	aatctataact	116580
tgcactacct	ccttcatattt	ccacaaatgt	ttgaagtgg	agaatttttta	aaaacttttga	116640
aggtacagct	gacagaattt	gctgtatgg	ttgaagtgg	ttgtatgaga	ggggaaaaaaa	116700

ggaataaagc	atgactgcat	ttttgtttg	tttgggtt	tgttttgag	acggagtctc	116760
actctcgcca	ggctggagtg	cagtggcgtg	atcttggctc	acggcaacct	ccgccttcgt	116820
ggttcaagcg	attccccctgc	ctcagcctcc	caagtagctg	ggactacagg	cgctcgccac	116880
cacgcctggc	taatttttt	ttttgttattt	tagtagaaac	ggggtttccac	cgtgttggcc	116940
aggatggct	ccatcteetg	acctccatgat	ctactcacct	ttggctccca	aagtgtcgag	117000
gttacaggca	tatataataag	catataaaatg	gtgttatagc	atacaaacag	gtatataatat	117060
aaacatgcag	tccacacacg	tgataggaat	gaggcagtag	tgaaggagaa	gttgatgttag	117120
gagaggggac	agttgttaca	ggaagaagaat	ctggaggccg	aaaggatgat	ttccagtgtct	117180
cacataga	attgtttaga	tggggacaag	gacaatttat	ctagagtccac	aggaaaagaat	117240
gcagtagacg	ggtagagatg	cagggtgat	gaaagatgtg	agagatgatg	gaaataaattt	117300
tctgattgt	tctatattct	caaggaagca	ggaagcaaaag	tcctcagcaa	agagaataga	117360
agaggtgtt	aatatttgag	aaaggagatg	tactgttagaa	aaaaaaaaaa	ctcagtttct	117420
cctctgtac	tctccacaaaa	cagaaccctt	ccatgactct	agttgtgtgg	ggttttttcc	117480
ctgtcagta	ccaatttgc	agatgttgt	tcagtgaaca	ccaaactgggt	gtctctaaag	117540
tcagttcagt	tctcacactg	tttacctgg	gatagcatca	gatcccacag	atttggact	117600
ctgtcccaca	agactgcctc	cacttcagat	gccagtc	agtacaagtt	gtggcctgtg	117660
cttctgactg	accttctata	aattggagtt	cccacagttc	cctcccttggg	ttcaataaaat	117720
ttgctagac	agctctcaga	actcaggaa	atgtcttaca	tatatttacc	catttattat	117780
aaaggatatt	acaaaggat	cagattgaac	aggcagatgg	aagagatgca	ttggcaagggt	117840
atggggagg	ggcacaagac	tttccatgcac	tcccaagggtc	atgcccaccc	ccaaagaaacct	117900
ctacagattt	agctatttcg	aagccccct	ccccattctg	tccttttggg	ttttttgtgg	117960
agacttcatt	atataaggat	gattgatcat	ttggctattgg	tgatcagctc	aacccctcagc	118020
cccccctatcc	cgggagggtt	gtgggttaggg	ctgaaagttcc	caaacgtgt	atctgcctt	118080
ggtcttctg	gtgtttagcc	tcatactcaa	agcttttag	aggccacacg	cacaagtcat	118140
ctcattagcc	ttcaaaaagaa	tccagagatt	ccatgatattt	taggcgtgt	atgctaagaa	118200
actggctaaa	ggccagttgc	aatgtctcag	gcctgtatcc	ccagcacttt	ggggaggctga	118260
ggcaggagga	tcgtttcagg	ccatgagatc	aaaaccagcc	ttgtcaacat	agtgagaccc	118320
ccttacaaaa	aattttaaaa	ttggccaggc	gtaatagctc	ttgtctgtag	tctcagctac	118380
tcagaaggct	gaggatcaat	gaggcctgg	tttggaaagca	gcagtggcc	atgatcgtgc	118440
cactgactcc	ggcttgggt	acaaatggag	acatgtctc	agaagaaaa	ggaaaaaaa	118500
aaaatgggc	aaaagactaaa	taacatattt	cacagtatca	cagattgtt	ttgtcttagga	118560
aagtgaatgt	aaacagacca	ggacactagt	atgatccctt	gttttcatga	aggtcccaact	118620
aaagtcatga	acacaaatgt	agactaggca	tcatgttata	ttgttttcc	agccatgttt	118680
aacagctac	taaataagat	attgttgc	ttcagtttat	tttagcgtt	ccttatttttta	118740
gcacatttca	tgttttaaaa	tttctaccaa	taacattttt	ataaactttt	tttacagataa	118800
cttcacaaat	ccataatattt	ttaagttaca	atcccagaaa	tagaattgt	cattgaaagg	118860
gtatgttcat	ttttaaaatgt	atgtagaaaa	ctgccaattt	gccttcagaa	aaaggtgttt	118920
gtatccccac	taacactatg	tttagttttc	ttgtgtccct	gctcaagtat	acatattttt	118980
aaaaacaatg	tttggccagt	tttactagata	aaaggtgtag	tgcccttctt	tttcaatctt	119040
tttgattact	agtggatgt	tatgttttt	cacgtgggt	atttttagtt	tttccctttt	119100
tggattgtca	tgtcccttgc	tcattttttt	tttggaaacat	ttcttagtag	tttataagag	119160
ctttggat	ttaaatgata	gtaacctttt	aactgtcatg	catgtcgaa	atcttttttc	119220
tgttggttt	cctttgtatt	tttgggggg	agggtttctt	tgtatagaa	ttaaatttttta	119280
tgttggtaa	tctttgtt	cattatgt	tcataatgt	tcaaaagact	ttcttatttttta	119340
agatcaatgt	ttacatgtt	tttctttttag	tttctatttttta	tttctttttag	tttctttttag	119400
tgtgtgtt	actcccaatg	tgatttcacaa	gtgtgtatata	atagttgaa	tttagtggca	119460
attaattat	ttacaacttc	tttgcagca	aggattttgt	gagaagatgg	acaggtggat	119520
cccaactgtt	tcgttttggc	acagttccata	gtcttttagt	caatggagca	agagaatgtt	119580
agtcattat	ttcacatgt	gtatgtttag	tttgggtt	tcatttttag	gaatgggtt	119640
cactcagta	aaaacaaaatg	ttttttttag	tttgggtt	tttgggtt	tttgggtt	119700
tataaaggat	actctatcc	ggtttacagt	tttgggtt	tttgggtt	tttgggtt	119760
aaaattttca	ttttcatattt	caatgttaat	atttccataag	ttaaaataat	tttgggtt	119820
ttcgtaaata	atttggggag	tgattgtatc	tgtatgtatt	atgacttata	gaattgggtt	119880
atttatttaa	ataatgtat	tcttcagatg	gtcttcttca	tttgggtttagt	aggcattttaaag	119940
cttaatggat	gtcttataac	taatccaca	tagattttt	gaaatggctc	cagagggttt	120000
ttagatttt	tactctatg	tgccctttaaa	aaaaatctat	tcatttttcc	acttaacatt	120060
tatcagaaga	gtgtctgtt	taagacgtgg	ttaggcatag	tgccagttt	gaaggaatgtt	120120
acagcctaat	aaaagacata	gggcattttt	tttgggttact	gtaatatgaa	gtggcatgtt	120180
ttaaatgtca	ggggagaact	acaaatgtat	aaaaagggtgg	gagagattac	atacaggtaa	120240
aggaatcagg	aatgacacca	tggggagtaa	ggtgttgg	acctaggcc	ttaagataca	120300

atagggacag	tatggaaaga	gtatattttt	cccacttaaa	ctcttcctt	ggtcgttccc	120360
tcaaatttc	ccttttgtcc	atgtgcaggc	actttagtga	gtttctgcga	agtcaccatt	120420
tctgaaata	ccagattgaa	gtgcgtacca	gtggaactgt	ttacctggct	gacattctct	120480
tctgtgagtc	agccctctt	tatttctctg	aggttaaagt	tgcatcttctt	ttcacactct	120540
attcgagcat	tccagcctt	aactatcaat	gctggggccc	tgtctatagg	aaataaacaca	120600
gaagagccaa	gtcatttcca	aaaagatgta	tcatttgc	aagtgttgc	tgatggcaag	120660
agtaatttaa	taatatatta	gagagaacat	gaaaatttca	tgtattaaat	aactctaatt	120720
ttgagaaacc	taattaaact	actgcatgta	agagagtgc	tgtttttat	tatttggagc	120780
tatTTTaaa	ccacagaat	tgaaatctgc	tccctgtca	taaatttcgag	accagacttc	120840
agaagagaaa	aaaatgtgt	aatttttct	tatgtctatc	attttactt	tagtcacttg	120900
ataggattgc	ccagtgaaga	agcatttgca	acagacaatg	agtatattaa	tctttttgag	120960
gcatacagtt	tagtataatg	ctcttgc	ggcttcaaca	agtgaatta	ttttgttggaa	121020
aagcaaatga	ctttaatgt	gaaagaggat	tcccagtctc	acaaggact	attttagaca	121080
ctcgattctg	cctcttaca	agaatacagg	tactcgttg	attttttctc	tcactccctt	121140
tcttgcttat	aagtttaat	caacaatttg	tttaggttta	tatgtctca	tgaatggtg	121200
gaaatgatca	gatataaaaat	atttggtttgc	tttagtttac	tctttatatg	tttgcgtggca	121260
aggaaccaca	aatccagttt	agtataattt	ttactctag	tcactaaaag	tttgcattcca	121320
gctgtgtagg	tagtgc	ttcttgc	tttttttgc	gtctaaaaga	actttaaaa	121380
actttcaat	ctcaatgac	tgtacttgc	tgacagggt	taacagaaga	agtagatctt	121440
tttgc	gttatttgc	tgtattttaa	tatttgcgt	tatagattag	agattgtgag	121500
agaaatctgt	ttatagtc	atttccctt	gtgtat	tcttcctagt	acatggaaaa	121560
agaggatgca	gtgaatatct	tacaattctg	gttggcagca	gataacttcc	agtcctcagct	121620
tgctgccaaa	aaggccaaat	atgtggaca	ggaggcagac	aatgtgc	tgtattttata	121680
tgacaagtga	gttattatg	tagatggat	cagcagat	ttatttgc	tttgcgtatgt	121740
tttgc	tttgc	tttgc	tttgc	tttgc	tttgc	121800
tttgc	tttgc	tttgc	tttgc	tttgc	tttgc	121860
cataaaagct	gttttat	gtgttgc	tttttttgc	tttttttgc	tttttttgc	121920
ctgctggaa	aggggggt	tatgtgggt	tttttttgc	tttttttgc	tttttttgc	121980
gtgttccca	gaagagagat	tttgc	tttttttgc	tttttttgc	tttttttgc	122040
agagaaggca	gaaaacaa	tttgc	tttttttgc	tttttttgc	tttttttgc	122100
ggggaaatg	gcacttca	gaaacttgc	tttgc	tttgc	tttgc	122160
atgaggatgc	tgtacta	tggaaatag	tgt	tgt	tgt	122220
tatattatac	tttctgtata	aatctgtca	ggcacgtt	taatttgc	tattttatgt	122280
ttcaactgaaa	atgaggat	ggaacatca	taacttgc	aaaatggaa	atatctggc	122340
aggcagatga	tgagcttg	gccagcttgc	taacttgc	tatttttgc	tttgc	122400
tcttactctg	tttttttgc	tttttttgc	tttttttgc	tttttttgc	tttttttgc	122460
acacttttag	tttttttgc	tttttttgc	tttttttgc	tttttttgc	tttttttgc	122520
gcaacatggc	aaaacccgc	tttactaaa	atacaaaat	tagtgc	tgatggcgt	122580
cgcctgtgt	cctagctact	tagggcctg	aggcagaagg	atcacctg	ccttgggagg	122640
tcgaggctgc	agtgcgtgt	gatccaccc	gggcaggc	tttgc	tttgc	122700
accctgtctc	aaaaaaaa	aaaacaa	aaaggat	tttgc	tttgc	122760
aatgctaaag	gggttaactt	gggatagaga	aaagtccaca	tttgc	tttgc	122820
ctaatagtat	ctaggccat	gggttctgaa	cattatgt	tttgc	tttgc	122880
tgcataaggaa	tcacctgaga	gttttattaa	aatagg	tttgc	tttgc	122940
acgcctataa	tcccagact	tttgc	tttgc	tttgc	tttgc	123000
tcaagaccag	cctggccaa	atgtttaaa	cccttctca	tttgc	tttgc	123060
caggcatgat	ggcacacacc	tttgc	tttgc	tttgc	tttgc	123120
ctcgagcccc	ggaggtggag	tttgc	tttgc	tttgc	tttgc	123180
gctgacagag	ggagactctg	tctcagaaa	aaaaaaa	tttgc	tttgc	123240
ccggcgttcc	acacctgtaa	tcccagact	tttgc	tttgc	tttgc	123300
ggtcaggagt	tttgc	tttgc	tttgc	tttgc	tttgc	123360
caaaaaat	actggcatt	tttgc	tttgc	tttgc	tttgc	123420
gcaggagaat	tgttgaacc	tttgc	tttgc	tttgc	tttgc	123480
aaaggttcc	atgtttctg	tctcagaaa	tctgatt	tttgc	tttgc	123540
gaatctttat	tttgc	tttgc	tttgc	tttgc	tttgc	123600
agtctaattt	tccttatttgc	caagat	tttgc	tttgc	tttgc	123660
caaagtca	aaacatgtt	tttgc	tttgc	tttgc	tttgc	123720
tccagtgttt	attctgttac	atgttgc	tttgc	tttgc	tttgc	123780
ataagatgttgc	agataggaag	tttgc	tttgc	tttgc	tttgc	123840
caagtgcgtgg	aggagtatttgc	accacat	tcaata	tttgc	tttgc	123900
tacaaaat	ctgat	tttgc	tttgc	tttgc	tttgc	

atggagtctt	gctctgtcac	ccaggctaga	gtgcagtgac	acgatctcg	ctcaactgcaa	123960
cctccacctc	ctgggttcaa	gcgggtctcc	tgcctcagcc	tcctgagtag	ctgggattac	124020
aggggggc	caccaccccc	agctaatttt	tgtatTTTta	gtagagacag	ggttccacca	124080
tgttggccag	gcccgtctg	aactcctgac	ctcaagtgtat	ccaccacca	ctgcctccca	124140
aagtgtggg	attataggcg	tgagccactg	tgcccagcc	gcttgtttt	gtatcatata	124200
tatgcatcat	cataatcatg	cattatcaac	ctttgtattt	ctgtcaggac	atagaaacca	124260
ttagagtgt	tggaagagag	cctttttttt	tttctcgcat	ttaatgtttt	tttgggtatt	124320
catttctaaa	tcaagttacc	aaaacattac	ctgcattata	ccccatcaag	gtagaaaatct	124380
tttgttata	aatattgggt	actcccttc	cacaccgagt	catcagtaag	tcctgttcta	124440
tccaaatagg	tcatatgcat	ctagctcacc	cctcagtgct	gttttgggtt	gaatttgtac	124500
atgtttactc	ctgatgcctt	gtagttatga	tgatgtgtt	ttatTTTttt	ctgtgcatac	124560
aagttctcg	ctcgctttt	agggaaaatg	accatgtctt	cctttctat	aaattccctt	124620
ctatctatca	agtcctcaac	agagaatagg	tacccataaaa	tatgtgattt	ttagtttctt	124680
tgcttcagtt	gtagtcgtat	ccttacagct	tttaaacaac	agtagagttc	accgtcaaga	124740
actaaggatg	gttggcaggc	agatagaaaag	gtagcaagg	gacccaacta	tctctgggaa	124800
agtggaaaca	aagaaagggtt	acatcaggcac	tgtcatcaca	tagctctata	gttctaggcc	124860
tgcaggctca	atcaagtagc	cttgatataag	attctctgg	ggaggtgctg	aaagtttgctt	124920
atacttgc	tggaaTTTGA	ttttatctcg	gatatTTTT	tccatagg	acttctccct	124980
ccaaagccaca	cattcttgc	gatttgcata	ttttgtacga	tttagaaattt	aatccaaatat	125040
ctgcagggaa	ggggggccac	tccccaactg	tttcacaact	ccattacgtc	aggcctggac	125100
aaccatggag	aaggtaaccc	agaacttcaa	acgttatcaa	ctacaagaag	ttttattgggt	125160
agaactcata	aaatataagg	tggggaaaac	aagcagaata	gcacagtgg	aattgaagca	125220
gtcagcggaa	gtgattaaga	gcagaggcct	tgagtctgg	ctggatgt	cagtacgtg	125280
ccacataaca	ttttatgtca	caggtagct	cgtgtacgt	ggtcctgtac	gattataatg	125340
gatcaaagct	ggtagtgc	taataacaaa	agttagaaaa	aataaatttt	aataagtaaa	125400
aaagaaaaaa	aaaaaaactaa	aaagataaaa	gaataaccaa	gaacaaaaca	aaaaaaaatta	125460
taatggagct	gaaaaatctc	tggtgcctca	tatTTTactgt	actatactt	taatcattat	125520
tttagatgtc	tccttctact	tactaagaaa	acagttaact	gtaaaacagc	ttcagacagg	125580
tccttcagga	gggttccaga	aggaggatt	gttatcaaaag	gagatgacgg	ctccatgcgt	125640
gttaactgccc	ctgaagacct	tccagtggtt	caagatgtgg	aggtgaaaga	agtggttatt	125700
gatgatccctg	accctgtgt	ggcttaggt	aatgtgggt	tttgttcttag	tttttaacaa	125760
acaaatttaa	aaagaaaaaa	aaaattaaaa	atagaaaaaa	gcttataaaa	taaggatata	125820
atgaaaatat	ttttgtacag	ctgtatgt	tttgttttta	agctgttat	acaacagagt	125880
caaaaagcta	aaaaaaatgg	aaacatttaa	aaatTTTact	aaatTTTattt	attattaaag	125940
aaaaaaattt	taaataaatt	tagttagcc	taagtgtaca	gtgtaagtct	acagtagtgt	126000
acaataatgt	gctaggcctt	cacattca	taccactcac	tcgctgactc	acccagagca	126060
acttccagtc	ttgcaagctc	cattcatgg	aaagtgcctt	tacagatgt	ccattttttt	126120
tcttttatac	tgtatTTTta	ctgtgcctt	tctgtattt	tttgttttata	cacaatttct	126180
taccattgtca	atagtggcc	acgatattca	ttatgtacat	atgtgtatca	gttttgcgt	126240
ccaaaagcaa	tagttgtac	catatagcc	agggggtgt	taggcatac	catctagtt	126300
tgtataagta	cactctgtga	tgttagcaca	atggcaagca	gccttaacgg	aattctgttt	126360
attgattgtat	tgattgtat	attgtattgt	acagaggttt	actccattgt	ccaggctgg	126420
gtcagcttgc	acagtctgg	cacactgca	tttgcctt	ccagggttca	ccaattatcc	126480
tgccctatcc	tcccaagtag	ttgggattac	aggcaggcac	caccatacct	ggctaatttt	126540
tgtatTTTtag	tagagacagg	gtttcaccat	tttggccagg	ctgttctcg	actcctgacc	126600
ttaagtgtatc	tgccctgttt	ggccctccgaa	agtgtgg	ttacaggcat	gagcttaccat	126660
gcttggccag	taactgaaat	tctctaatgc	cattttccct	atctgtaaag	tgacgataat	126720
atgcacgtt	acctcaaaat	tactttgtat	attaaatgtat	ggtaatgtat	ataaaatata	126780
tattaacata	gtacctgaca	catgttgc	atcaaaaaat	tttactact	tttattacta	126840
ttattattac	gtatTTTtaa	aaatattagag	agcagtatca	aaaatttagct	gggcgttagtg	126900
gcatgcacct	atagttccag	ctactcagga	ggctgaagct	ggaggattgc	atgagcctgg	126960
gaattaaagg	ctgcagttag	ccgtgttcat	gcccctgcac	tccagcttgc	gtgacagagc	127020
aagaccctgt	cttgaacaaat	taaagaaggc	attatgcgc	aacgttagt	tagaaatgtat	127080
ccacatataat	caccgttaac	tgtcaacagg	atggaaaccc	tagtttggg	tattatgtatc	127140
acaaggatatt	attaatagct	tattaataat	aaagcggttg	ctaggcacgg	cgactcacat	127200
ctgtatccc	agcacttgg	gaggccgagg	tgggtggatc	acctgagg	aggagtttga	127260
gaccagctg	accaacatgg	agaaaccccc	tctctactaa	aaataca	ttagccgggc	127320
gtgggtgtgc	atgcctgtaa	tcccagctac	tttagggagct	gaggcaggaa	aatcttgg	127380
accggggagg	cagaggttgc	agttagctga	gatgcacca	ttgcactcca	gcctgggcaa	127440
caagagcaaa	actccgtctc	aaaaatataaa	ttataataaa	taaataaaaag	taaagtattt	127500

atgtttgtga atgatttatt cttctaata	actagaggag	atttttccag	gaatttcaga	127560
gccagtggg ttatgtgtc	tgtatgtgc	atgtgtatcc	aggtgaaaaa	127620
acgttatt ataaccat	acataaaaac	tgaattttag	gaatactgaa	127680
tagaagtcaa atcattaaat	agcttagtgc	aaacagaata	gagtgtcagc	127740
tgtatgataat	attttcacga	ttaaaattaa	acctttctg	127800
atctgtatca tataaagaat	gtaaaatttc	aggtaataa	aattaaaatg	127860
aatgcaaaaa tagttcttac	tagatgtgt	tatgtaaagg	acttagacta	127920
cactgtcaag accctggtag	ttaggttaga	aaaagagat	gaatgatca	127980
actttgatgtttctgt	atgtgtgc	agatgtgt	gttacagtgg	128040
ctgctatcct ggagcttagt	ctacaaaaaa	ggtacatatt	ggccgggcac	128100
gcctgtatac cttagcactt	ggaagatcg	ggcgggtgga	tcacctgagg	128160
aagaccagct tggccaacat	ggcgaaccc	cgtctctact	aaaatataca	128220
ggtgtgggg cggcacac	taatccac	tactcgggag	gctgaggcag	128280
tgaacatggg agacagagg	tccagtgc	cgagatcat	ccactgcatt	128340
ggacaaaagc gaaaatacgt	ctcaaaaaaa	caaaaacaaa	caacaaaggc	128400
tacgaacata aatatttaca	aattatactg	aataagtct	catgttttat	128460
ccaggatcaa acttttccct	cgttagatta	gaatataaaa	taataaaccat	128520
tcagtataat taataattat	ttaatgtaaa	taaaaacatc	tatgtacaat	128580
ttaagaattt atttgaaaaa	aaaacaatgt	ggaaacagat	attttgatat	128640
attgaaattt ataatgttct	tttgaagagt	aaagtgcacca	tatataattaa	128700
taactcagca atcacacccc	ttgtgagtt	tcttaaggaa	atcagttga	128760
aatatatgca caaaagactt	aacatttac	ataaaaccaga	aaaatcgagt	128820
atccatgtca ctatttctg	ctaaaaagta	ttaatataat	ctttatgtaa	128880
acaatattt tggggggagaa	aaaccaacaa	aattacatgc	attgtatatt	128940
tttttttta gacagtctt	ctccagcgtc	caggctggag	tgcaagtggg	129000
tcactgcaac ctccatctcc	cagggtcaag	caattcttct	gcctcaggcc	129060
ctgggattac aggcgctcac	caccatgcct	actaattttt	tatgtttt	129120
gggtttcatc atgttggca	ggctggctt	gaactctgg	tctcaagtga	129180
tcggcctctt agagtgcgt	gattacagg	gtaaagccact	gcacccagcc	129240
taattttat ttgtaaaactg	tacaaaggga	taataacttgc	agtacaacaa	129300
cattgttat aggtgtttaa	catttgaac	cagtagaaatt	ataggtaaaa	129360
taaaacagtt ttatgttgat	ttgatttca	ctttaaaaata	atgctttca	129420
gtcttttgc ctggctttt	gtccagcaat	cttttataa	aatatttggaa	129480
cattcggttc gaggagat	attttggc	ggaaacgtgt	cgctgactgc	129540
gttggccctc ctgtgatgtc	tcacccagg	agttctgaca	gctctcggtc	129600
actgattgcg tctgcccatt	gggagaaaag	catacacatc	ctttccttca	129660
acagatccca ttatgtttaa	atttttaagtt	gtggaaaaaa	aagataaaaag	129720
tggccctgtgc ctgtatccc	agcacttttg	gaggctgcgg	ttggcgatc	129780
aggatctega gaccgcctt	gcccacatgg	tgaaacccca	acacgaggc	129840
attagccggg catggggca	ggcacctgta	atccatgtca	cttgggaggc	129900
gaatcgctt aacccaggag	gcagagg	caatgaacca	aaatcactcc	129960
agcctgggtg acaaagttag	actgtgtctc	aaaaaaaaaa	aaaaaaagaga	130020
tagcttactt acatcttct	aatcaagca	tttgggtgtaa	ctttaaaat	130080
aaagatctatcat gtgtttcat	ttagggcatt	attttatgtt	aatctgtggc	130140
aataaaatcaa gtaatatgaa	atattatcat	agcctctgaa	gagctcttta	130200
tatttaggtt actttttgt	aaataaagtga	atgaatttctt	aggtcttctt	130260
ttcttgagac agggtcttct	cgctgcaacc	ttggaaattt	gggctcaat	130320
ccacagccctc ctgtatgt	gggactagag	gcatgcacca	ccacgcctgg	130380
atttttttt ggcaggcat	gatgggtcac	gctgttaatc	ccagcacttt	130440
ggcaggcaga tcacgggtc	ggggagatgg	gaccagccctg	gccaacgtgg	130500
tctctactaa aaataaaaaa	attagctgg	tatgtgtggct	catgcctgt	130560
cttgggaggc tgaggcagg	gaatggctt	aaccaggggag	tcggagggtt	130620
agatcacgcc actgcactt	tgcatgtga	cagagtggaa	ctccatctca	130680
tttttttaaa tgatgggtc	ttgtgtgtt	gctcaggctg	gttcttgc	130740
aatgcgcct gcttcagctt	aagtttctt	ttttttgtt	aaagagacagg	130800
gttggccagg gtgtctcaa	actctggct	tcaaggcagtc	ctcccaac	130860
agtgtggaa ttacaggcgt	gaaccactac	ctataatgtt	gtgtttact	130920
tgatttcggtt ttgcattacc	gtgcacatt	gtgcatttcc	ttgaccc	130980
tggagtgcatt tcatatgtt	aaccatacct	gatttccctc	aaaatcacac	131040
atccataagac aagaaatcta	aggaggcata	aagaagttaa	ctggtttat	131100

98/122

cagtaaatga	tagagccaga	aatattcccc	ttctagtgtt	cttcaccatc	agcttaatgt	131160
agcataataa	ttttctaatt	actgttgaca	aataaataac	cctttgaatt	ttcaataactg	131220
ggccttggat	aaatttcct	aattttaag	agagtataag	cgtattgcca	tttacaaaggc	131280
tctcttgagt	atctttttt	tctgttaagt	ttaaccttagga	gataaaactgc	tgagtatgtt	131340
tgccatTTTg	gtttttgtat	ataggtaga	atgtcttggg	tttttttttt	tttttttttg	131400
gtttttgttg	ttgtcattgt	ttgagacagc	atcttgcct	gtcgcccagg	ctggagtgca	131460
atggcacat	cgtggctcac	tgcaacccctc	acctcccggg	ttcaagcaat	tctcctgcct	131520
cagcttccTG	agtagctggg	attacaggca	tgtgcaacca	cacctggcta	attttgggt	131580
ttttagtaga	gaagggggtt	caccatgtt	gtcaggctgg	tattgaactg	ctgacactcat	131640
gatccacTg	cctcggcctc	ccaaaatgtct	gggattgcag	gcatgagcca	ctgcacctgg	131700
ctgaatgtct	ttttttgtat	taggcactta	agaaaaggcc	aggtactaac	cataaaatat	131760
attttatac	cttttggta	tactatata	atagaaaact	gcacttatca	taacctttaga	131820
cacccTGAAG	aatgttcaca	agcagaata	accatgtga	cccagcatcc	agatcaaaaa	131880
cagcattatc	agccccctct	gaagccctt	tgggccctt	ccatttactg	tcttcttgc	131940
caccaggta	gctactatcc	tgactttga	ttggcatagat	tagcattacc	tgttcttgc	132000
attttataaa	taaaaccata	ctgttatttc	ttttcttgc	cagcttattt	gtgctaattc	132060
acatttacat	catacaattc	agtggTTTT	atatggtcac	agagtttagt	aaccattacc	132120
acatcgattt	tagaacatTT	tttttactcc	agatagaaaac	ccccttactt	taactccaa	132180
atccccact	ccaccagccc	taggcagcca	ctagtactat	ttttatctct	atagagacaa	132240
tagatttgc	tattctggac	atttccataaa	catggaaacc	tatattatgt	ggtctttgt	132300
tgccaactgt	ttttcaactt	gcacatgttgc	ttcaaaagag	catcatgtt	tccatgtttg	132360
gcatgtatca	gaattttatt	cctcattatg	gccaaatatc	ccattgcaag	gatttatgac	132420
attttatttg	aattgtaccc	ttcttctgc	cattttatcaa	taatgtact	tgaccattt	132480
gttacaagt	ttttgtgtgg	atacaggTTT	ttttttgtt	tttttttttg	agttggagtc	132540
ttgtctgtc	gcccaggctg	gagtgcagt	gcacaatctc	ggctcaactgc	aacctctgtc	132600
tcctgggttc	aaggacttct	cctgcctcag	cctcccgagt	atctggact	ataggcacgc	132660
accaccacgc	ccagctaatt	tttttagaga	gatggggTTT	caccatgtt	gccagctctgg	132720
tctcgaactc	ttgacctcaa	gtgatccacc	cacatcgggcc	tcccaaagt	ctgggattac	132780
agggtgagc	cactatggcc	ggctgtgtt	ttcatttttt	ttgttgtata	tacataggag	132840
tagaattgt	gagtcaagag	gtaaatctta	aacttttga	aaaactgc	gattgttttc	132900
cgaaaaggct	gcacatttt	gcaatcccac	cagcagtgt	tgagttttac	agttcttcca	132960
catttcattg	gaacttatta	tctgtttggc	tgttttttaaa	aatgatagtc	attccaataaa	133020
gttctacttc	agtgtggTTT	ttgcacttct	ctgatgagta	atgatgttga	gcatctttt	133080
atttgcatt	tggcTTTGT	tctactttt	gaaaaatgtt	tatttacatc	tttggccat	133140
tttttatttt	attttttttt	tttttgagacc	aaagtctact	ttgtcagcca	ttgtcagcca	133200
ggctggagta	caatgggtgt	gtctcagetc	actgcaccc	ccgccttctg	tggtcaagtg	133260
atttctctgc	ctcagccccc	cgagtagctg	ggattacatt	tcaggcacct	gccagcatgc	133320
cggcgtattt	tttgatTTT	tactagtgac	agggtttcac	catgttagcc	aggctgggtca	133380
caaactccTg	accttcagggt	atctgcgtc	ctaggcttcc	caaagtgtc	ggattacagg	133440
cgtgagccat	tggcccage	ctagattttt	tttttttttt	tttttttttg	gaaggaggct	133500
tgctcttgc	gcccaggctg	gagtgcata	gcacaatctt	ggctcaactgc	aacctctgtc	133560
tcctgggttc	aaggcatttt	cctgcctcag	cctccccagt	agctggatt	acaggtgcct	133620
accaccacac	ccagctaact	tttgatTTT	tttttagagac	agggtttcac	catgttggcc	133680
aggctggTc	caactccTg	cctcagggt	tccacccgt	ttggcttccc	gaagtgttgc	133740
gattaccggc	atgagctacc	agggccagcc	attttctca	ttatattgtcc	caggctggtc	133800
tcaaactcct	gggttcaagt	gttccctctg	ccttggcttc	ccaaagtgt	gggagttacag	133860
gctgagcca	ccttgctcag	ccccTTTggc	cattttttaaa	ttagattgtcc	ttttttatatt	133920
gagtttcagg	agtccTTT	atattctaga	taaatgtccc	ttatcaaatt	atattatttc	133980
cagttatTTT	tttcattctg	tgaggTTT	ttctcttacc	tttttttttt	gttgggtttt	134040
tgttttttgt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	134100
gcagtggcac	aatcacagct	cactgcctcc	tcaacttct	tttttttttt	tttttttttt	134160
ttcagccctc	ctgaatagct	agggccatag	atacacacta	tcacacccag	tttttttttt	134220
ctgtttgtag	agacagatct	tactgtttt	cccaaggTTT	tctcaactc	taggtcaaa	134280
gtgattctcc	cacccTCTG	tcccaggagt	ctgggattac	aggtgtgagc	cacacgcAAC	134340
ctgtcttttc	actattataa	gtgtttctt	cttgcagct	cccgagtagc	ttggattaca	134400
ggoacccacc	accatgcctg	gttaattttt	ttgcattttt	agtagagaca	gtgtttcac	134460
atgttcaccc	ggctgggttt	gaactccTg	cctcagggt	ttcacctgtcc	atggcctccc	134520
aaagtgtgg	gattacagge	gtgagccact	gcacccggcc	aaaatattgc	tttcttaaca	134580
gtattgtctt	ctaatttttg	aacatggat	tatcttcatg	tatTTATGt	tttttttattt	134640
tcagcagaat	tttgatTTT	tcagatgaga	agcctttcac	ctccttgggt	cattttattcc	134700

99/122

tatgttttaa	gttcttttcg	attccattat	aaatagaatt	gttttcttaa	tttcattttc	134760
agattgttg	atgagagagc	atagaatac	aagtgtttt	tacatgttga	tcttgcact	134820
tcaacttta	taaatctgat	tgttagctct	aatagtttc	tttgtggattc	tttaggattt	134880
tcaatata	agatcatgtc	atttatggat	agagatagg	ttttttctgg	ctagaactta	134940
cagacaat	atgagtagaa	gtggcagaag	aaaaaatctt	tgtcttgtt	cctatctgac	135000
agggaaagct	ttcagttca	tcatttata	tgatgttagg	tgtggggttt	caataaatgc	135060
cttttttcag	attcaggaat	ttccctatca	ttccgtattt	ttaaggctt	tttttttttt	135120
ttaaatcatg	aaagggtgtt	gaatattgtc	atgttcttc	tgtatcagta	taaatgatcc	135180
tatggattt	gggttttatt	ctgttgatgt	gaatattaa	ttgattttca	gatgttaaac	135240
caacatttgc	tacctggat	gaatctact	tggicatggt	gtataatctt	ttcaatatgc	135300
tgctggattc	catttactgg	tattttgtt	aagattttgt	atctgaacgc	ttaagataac	135360
atttacactc	tatcagaaat	gaattgacca	taaatgtgag	agtgtatttt	tgggttcttg	135420
attctcttcc	attccaaaga	tagacataca	tccgtctgt	tgtctgtt	tatgccagta	135480
ccatacttc	ttgattacta	ttgcttgtt	ataaagtttt	aaatcagaaa	gtataaatga	135540
gattttgtt	tctggatcac	agtcttctata	gaatttagtt	gaaatattt	cctcttattt	135600
ctggccccc	tttttttttt	gtttaactgt	gtatcttgg	gatgttctt	tctcaacaca	135660
ttagagccgc	tttccttacc	ctccccacccc	tgtatagag	aggtctataa	gtgtctgtt	135720
aattattttt	tttacttaac	ctattactta	gtcggggaca	ttaagcttgt	ttatgtcttt	135780
tatTTTAAAC	aatgtcgacg	tgaataatctt	tgtatataag	tcattttcca	tcaatataag	135840
tctctctgt	actgaaatttt	tagaagtgg	attttaggt	caacctatgg	ctctgttattt	135900
cacaaaaaaa	ccaaattctgg	tttttttttt	ggagggtgggg	agttaggagg	agaatgtctgg	135960
aggagaactt	gctgtactca	gctggctagt	cattttagaa	aggtttctt	agcttctttt	136020
tgtcatatgg	cctcaccctt	aatcaaaaaac	attcttattt	accctgtaaa	catggggctt	136080
tactaccctt	gatacatatt	tctggatgt	tgacagcttt	tcattattgaa	gaaataatgc	136140
tgtgagttaca	gcacattttgt	ttggaaacttag	gtcgtaaga	atgtctttt	aattcataca	136200
ttatacattt	tatttttttt	tatttttttt	ttttttttttt	agagtcttcc	tctgtcgccc	136260
aggccagcgt	gcagtggatc	aatcttgggt	caactgcgacc	tccatcttct	gggctcaagt	136320
gattctcatg	tctcggcttc	cagagtagct	atgggtacag	gcatgcacca	ccatggccgg	136380
ctaatttttt	tatttttagt	agaaaactggg	tttcaccata	ttgaccatgc	ttggcctcgaa	136440
ctcttggctt	caagtgtatcg	gcctgcctca	gcctccaaaa	gtgctgggat	ccttgttattt	136500
ggtaaaaaagat	gaatattttag	ggctgtatgg	tggctcatac	ctgtatcc	agcacatttct	136560
gagactgttgc	ttggggaggat	cttggagcccc	aggaggggtga	gctgtcagt	agtgtgtatc	136620
gcccatttgc	acttttaccc	aggaattata	ggcttcagtc	actgtgcccc	gcatgtacat	136680
ttaatatttgc	tgctttcttc	tttttagctat	agtatgaggt	tacattttag	agtcatgtt	136740
gttaagcatc	ttaatagtga	tgaggtttag	tgaagtttac	ttcttatttca	aacactgaag	136800
aaaattttgt	acaaatctgt	cacattttttt	gtttttttttt	gattgtttca	tatacttca	136860
attttacaaat	ttcttatttga	gtccagtgt	aaaaaaagcca	gttatttttt	actgaaaaat	136920
tttggatggat	cgatattttgt	ggatggggca	agtctggatc	cagaatcttt	atatacaacgg	136980
acatatggcc	ggtaagctt	gctcatgtt	agaattttta	caagtgtaaa	taactttgca	137040
tcttttaaat	ttttttaattt	aatttttacat	ttttttctaa	tcttatttttta	tatgcccaga	137100
acttttacctt	agagtgttgc	gttataatgtt	gtgggtttaat	ataaaaggctc	tggagtgtact	137160
tcttgggttt	taatcttggc	tctggccattt	tttgggtttaat	gttatacttct	ttgttatctca	137220
gttttcttcat	ctgtaaaaatg	ggataatataa	agtggaaaaga	tttggccaaatc	atttactcttgc	137280
ggctgcataa	ctgtatcttg	aaaaaaagtat	tccttttagt	ttaagaatttta	agttgggtat	137340
tcatttttgc	ttgtataaaa	aagatagtgt	ttcataggat	atgcccactt	ctgaaatttta	137400
ccacagatcc	aatcataaaaa	tcacttttcc	ttccctttaaa	atagcttgc	taacatgtaa	137460
aggtgttgc	aggcttgcatt	acacttccct	gatccgtacc	ccatgtcccc	gcagcaccat	137520
aaaaaaaggat	tttcaacata	ttttaattact	ttcagtagaa	gttataacttct	ttgttatctca	137580
cgcagtgtct	cacacctgtt	atcccagcac	tttggggaggc	tttggggaggc	tttggggaggc	137640
gtcaggagat	tgagaccatc	ctggctaaaca	cgatgaaacc	cgatgtcttac	aaaaataaca	137700
aaaaatttagc	cgggcatgtt	ggcaggccacc	tgtgttccca	gttacttggg	aggctgagac	137760
aggagaatgg	cgtgagcccc	ggaggccggag	cttgcagtgt	gtttttttttt	ttttttttttt	137820
ctccggcttgc	cgcgttggat	cgagactttt	gtctaaaaaa	ttttttttttt	ttttttttttt	137880
ttggggagact	gaggggctta	aaaaagtactt	gaaggaaagta	ttttttttttt	ttttttttttt	137940
tgtattttgg	aaggccagtt	tttcagctgt	gtcagctttt	ttttttttttt	ttttttttttt	138000
cttttagaaa	ataacggaca	aggccggggca	cggttgcata	ttttttttttt	ttttttttttt	138060
tggggggccgc	agacggggccgg	attacctgtat	ctcaggatgt	ttttttttttt	ttttttttttt	138120
tggtaaacc	ccgtcttcata	aaaatataaa	aaatgttgc	ttttttttttt	ttttttttttt	138180
gtgtcccttgc	ctactccggta	ggctgagggtt	ggagaatttgc	tttgggggggg	tttgggggggg	138240
ttgcagtgttt	ccaaagatcac	accatttgcac	ttgcagcttgc	ttttttttttt	ttttttttttt	138300

100/122

ctcaaaaaat	aataataaaa	taaaaaagaa	tggacagtaa	acctaataatga	gttcatttccc	138360
aaagatgatg	ttattcttaa	gggatggttc	atttatttaa	gacccttacat	aaagtctatc	138420
aattgcgtga	tttttcactt	ctgttaattgt	gtgttatgtat	aatgtaaaata	tatatgtttt	138480
tgttttgttt	tggttttttg	agacggagtc	tcgctctgtt	gctcaggctg	gaatgcagtg	138540
gtgcaatctc	agctctctgc	aacctctgtc	tcccaggttc	aagcgttct	tctgcctcat	138600
cctcccaagt	agctgggact	acaggcacgt	gccaccacgc	ccggcttaatt	tttggatattt	138660
tttagtagaga	tgggggttca	ccgtttagc	caggatggtc	tcaatctct	gacctcggtga	138720
tccacccggc	ttggcttccc	aaagtgttgc	tattacaggc	atgagccacc	acacccagca	138780
tgtatTTTTT	aaatgtataa	aatgaagcag	aaaagagaaaa	tgataatTTT	tcttcatctt	138840
gaaagattat	cttcaccagg	cgcagtggct	cacacttgc	atcccagcac	tttgggaggc	138900
ctcggcaggc	ggctcaactg	agttcgaaac	cagcctggcc	gacatgggtga	aactccgtct	138960
ctactaaaaa	taaataaaa	aagatggtt	taatatatgt	tttagttta	tgatTTTtagc	139020
atcttctga	aattttctc	aaggcaagta	aatttgcattc	agttggata	tttggtaccca	139080
tctatgaaat	aacttattag	gaagataatct	ctaaaataag	atcactttgc	ctaaaataaaa	139140
ctgatataatt	gatgttcaca	gaattttct	ttaaccgac	ttgataaaatg	cattattctt	139200
gacgtcaagt	gatccaccc	cctcagcctc	ccaaagtgt	gggattacac	acatgagcc	139260
ccgcacccgg	cattattctt	ataaaagggt	aaatttctag	ttaagttta	tgcctcttt	139320
gttcatgtac	cattgtttat	tttcttccct	tcctactcac	agtaatcatt	cttatggtat	139380
gcactttgt	ttgcttattt	ttatgttaatt	gatattacgc	tccattctgt	acgttgtact	139440
ttcattcaca	gtgagttttg	gacattctta	tggtcatcta	tacagactta	cttcattttt	139500
actacactgt	agtattccgt	atgtatatt	tactataact	catcaactgt	gcagagcatc	139560
tcatagtgt	tgttattact	ttttgcatt	ttgttatca	tgagtattta	agtcatttgc	139620
agttttccc	tcttatacccc	agtattacag	aggatcttct	tttatatgt	tctttgtacc	139680
aaggaggcaga	ttaaaaaatt	ttttttgaa	aaaatttttt	aaaaaaaaatg	aatgaagtc	139740
tcactatgtt	gcccaggctg	gtctcaaact	cctaggctca	agcaatcctt	ccatcttggc	139800
ctcccaaagt	gctgggggtt	caggcatgag	ccaccatgcc	tggcttacat	tttaaattttt	139860
gatagcttt	acaatttact	ttgtaaagta	tctgcattcat	tttatgttct	caccagtctt	139920
taataagaat	acttcataact	tttgcttgc	cacagtggct	cacgcgtgt	atcccagcac	139980
tttgggaggc	cgaggcggggc	agatcaagag	atcgagacca	ccctggccaa	tatggtggaaa	140040
ccctgtctct	actaaaaata	caaaaattag	ctgggcgtgg	ttgcgcaccc	gtagtcccag	140100
ctactcgaga	ggctgagaca	ggagaatcac	ttgaacccgg	gaggtggagg	ttgcagtgaa	140160
cttagatcac	accactgcac	tccagcctag	caacagagtg	agactctgtc	tcaaaaaaaaaa	140220
aaaagaatac	ttcagactta	atttttttc	cagtcttaag	tgtttctaa	tgagatttgag	140280
tttcttttgg	tatgtctct	gatgttctag	tttttttctt	ttatgaattt	actgttcatc	140340
tcttttcc	attattttctg	ttgggtgatt	ttattatgt	tttgttaaaa	tctgtatata	140400
tttttcagca	tgacacttca	ttattcaaaa	aaaaaaaaag	attctctatg	tttctcgata	140460
ctaatacattg	gttggtaata	cctttttttt	aagaccctta	ctgttattttt	tgtttttttt	140520
ttttttttt	ttttttttt	tttgagatag	agtctgtct	tgttgcctt	gttgggagtgc	140580
aatggatgt	tctggctct	cagctact	caactgttac	cttacatcc	ctgtttcaag	140640
caattctct	gccttagctt	cccaagttagc	ttgggattaca	ggcatccacc	accacacccca	140700
gctaattttt	gtatttttag	tagagacagg	gtttcaccat	gttggccagg	ctggctctaa	140760
actactggcc	tcaagtgtatc	cgccctgcctc	ggcatcccaa	agactctgg	ttacaggcat	140820
gagccacagt	gcctagccac	ttttgtttt	ttaactttgt	tttatatgt	tatagttttt	140880
gtataaacag	atgtatgtat	acacacaact	atggctttat	aatatgtttc	agtcattttt	140940
agagcaaggc	ctaccccttt	gggtgttctt	ttacaaaatt	gtctttgtca	tttttgttgc	141000
tttttctta	tttgtgaatt	tttagattgt	gaatttacctg	ttgacttacc	atgtttgtt	141060
aactgaggat	tttgaatgg	attgcactca	attaaagatt	atcttgcattt	ctgtgcagca	141120
atgtttttt	tcaaaaataatc	cctactttaa	attacttagg	atagcttata	atgtgtttt	141180
tggtttctt	gatttagat	aaacgctttt	attgttgc	tttcttctaa	attaaaaact	141240
gattgtttaga	agttttaagg	ttctgttcat	tcttttttag	gaagatgaca	tttggaaagag	141300
tcagtactt	ggggcaattt	atccgagaat	ctgagcctgt	acctgtatgt	agggaaatcaa	141360
aagtttgc	gttggggat	acttcatatt	aagcctttac	tcacattagt	gattgactgt	141420
aagtcaaaaga	ccacttaagg	ttttaactgt	ttttttgtt	aagtaaccac	tgtatcttt	141480
accttgcgtt	tatagtccaga	agtaagtaca	agggtttccct	gtatgtccat	tttttgttgc	141540
tctctcttgc	atcaaaaatgt	agtgaacttg	cttgcctt	ccagaaggca	catgaatata	141600
aaaaaggcatt	gtcttatttt	ttatTTTat	gaaaaatacc	cgacctaagt	tggacttaat	141660
gtttgagacc	gttttttttta	ttaaattata	ttttttctt	tttctttttt	ttttttgaga	141720
cagttcttgc	tctgtcaccc	agaccggagt	gcagttggct	gaccgcaccc	cactgcaacc	141780
tctgtttctt	aggttcaagg	gattttctg	cctcatccct	ctgagtagct	gggactacaa	141840
gtgcgcacca	ccacacccctgg	ctaaatTTTt	tatTTTtagc	agagatgagg	tttcaccacac	141900

101/122

ttggctaggc	tggtctcata	ctccgtaccc	caagcaatcc	atccgccttg	gcttcccaa	141960
gtgctggat	tacaagtgtg	agccaccatg	cctggccta	ttaaattatt	tttattaaat	142020
ttctcaaga	ttgatgaaag	taatgaaaata	taaaagtaat	gaaatatatg	tggaaaatag	142080
actggattaa	gaaaatgtgg	cacatataca	ccatggatac	tatgcagcca	taaaaaagga	142140
ttagttcatg	tcctttag	ggacatggat	gaagctggaa	accatcattc	ttagcaaaact	142200
gtctcaagga	tagaaaacca	aacaccgcac	gtctcaact	ataggtggga	attgaacaat	142260
gagaacactt	ggacacaggg	tgggaaacat	cacacgctgg	ggcctgtcg	gggggtgggg	142320
gctggggag	aatatggatt	aggagatata	cctaataata	atgacgagtt	aatgggtgca	142380
gcacaccaac	atgtacatg	tatacatatg	taacaaagct	gcacgtgtg	cacatgtacc	142440
ctagaactt	aagtataata	aattaaaaaa	aaataaaat	atgtgaaaaa	tattaatagg	142500
tcaaattca	aatttgtcat	ttaatcagaa	gagtagttt	gtcaaatcca	agggttagac	142560
aacagaaatc	ttttttgtca	agtgcattct	ttgtgactgt	tttcattttc	ttcctgggtt	142620
acacagggag	atttcagaaa	caaatgtgg	tccgtgacag	atgttatcta	gaagttttta	142680
gttgggtga	atttgcagta	ttttttttag	taaaagatata	taatttttgt	aagaagaaaa	142740
attcaatttt	gataagtatg	ttaagatta	agagctattt	gccaggcgt	gtggctcatg	142800
cctgtatcc	tagcaattt	ggaagctgg	gcagggtgg	cacgaggctca	agagattgag	142860
accatctgg	ccaaacatgtt	gaaaccctgt	ctctactaaa	tttagccaggc	gtgggtggcac	142920
atgcctgtgc	acccgcctcc	gggtttaagg	gatccactt	cctcaggctc	ctgagtagct	142980
gggatattac	gcccgcattgc	taatttttgc	attttttagt	gagacagggt	ttcactacat	143040
tggccaggct	ggctggct	caaactcttg	acctcagggt	atctggccgc	tttagecctcc	143100
caaagtgtg	ggattacagg	catgattcac	catgtctggc	catttatctt	atttttttttt	143160
tttttttttt	ttttttgtga	gacggagttt	tgctgtgtcg	cccagagctg	gagtgcata	143220
gtgcgtatctc	agtcactgc	aacctctgcc	tcctgggtt	aagcaatttct	cctgcctca	143280
tcttcaagt	agctggattt	acaggcgtgt	gcccacat	ctagctaatt	tttgttatttt	143340
tagtagagac	agggtttcac	catgttggcc	aggctggct	cggaactctt	gacccgttaa	143400
tctgcccacc	tcggccctcc	aaagtgtga	gattacaagt	gtgagccact	gtgcccagcc	143460
atcttatttt	ctttctttt	ttttgtcggt	tggggggggg	acagactcta	gtctgtcg	143520
caggcttggc	tcactgtcaac	ctctggcccc	cagggtcttag	caattatttct	gcctcagcc	143580
ccaaagtagc	ttggattata	ggcacatggc	acacgcctt	gctaattttt	tgtttatttt	143640
agtagagatg	gggttttgc	atgttacca	tgctggccct	aagtgtatcc	cccacccctgg	143700
cctcccaaag	tactggctt	acaggcgtga	gcttgtattt	gtttaaagaa	caatatttgg	143760
ggetgcattgg	tggttcatac	ctgtatctg	agcactttgt	gagactgaga	tggaaaggag	143820
gttggagggcc	aggagggtgt	ggctcggtgt	gcagtgttt	gtgatcacgc	cattgcactt	143880
ccaccttagt	aatggggat	gaccatgtct	ctaaaaaaa	aaacacaaatt	ttttttagga	143940
atactggaa	gagggtcgat	gtggttttag	aacagagggaa	gtgcccagat	acctttgtga	144000
ggcattggcc	aggaagaact	ctacagtgtc	tttaggttag	ttctgtccat	aaggataatg	144060
gggtctccic	cccagtatta	atagaaaaatc	tctgagttgt	ttttttttgt	ttgttttttt	144120
tgtttttttt	tcctgatgt	gagttctct	ctgtcgccca	ggctgggtgt	ctgtggcgcc	144180
atcttttttt	actgcaagct	ctgccttcca	ggttccacac	attctctgc	ctcagcctcc	144240
caagtagctg	ggactacagg	tgtccaccac	cacgcccac	taattttttt	tttttttttt	144300
tagagatggg	gtttcacat	gtcagccagg	atggtctcga	tctctgtacc	tcgtgatccg	144360
ctcgccctctg	ccttgcaag	tgctggagtt	acaggcgtga	gccccatgtc	ctggccctgg	144420
ttttttgttg	tttttttattt	tttttttttt	tatttttttt	ttgagacaga	ctctcgctct	144480
gtcgcccggt	ctggaggtt	gtggcacat	gtcggtcac	tgcaagctt	gcctgcccagg	144540
ttcaagccat	tctcctgcct	caggctctgt	atgtatcc	tgtagtgcata	gtgttaacat	144600
cgcggccgtt	atttttgtt	tttttagaaag	agacgggggtt	tcaccgcatt	agccaggatg	144660
gtctcgatct	cctgtatgt	tgtatccccc	acctcggctt	cccaaagtgc	tgggattaca	144720
gggtgtggcc	accgtgcctt	ggctgtttt	tttttttttt	taatctggtc	tcatacctct	144780
gacagctcat	gaagaagtgc	tctctgttca	tatgtatat	tgtagtgcata	gtgttaacat	144840
agcataggtg	ttcgggtttt	gcagtttctg	tttggtttat	atgatataag	gtgttattatg	144900
agcagttgaa	gatataatagg	aaattttttc	ccaaaccact	atctctgttc	gttctattca	144960
ttcagtctgt	ttatgttatt	ccttcattca	ttcattttat	agaacagtgg	agtgcctact	145020
gtatgcattt	attgttctgg	gtctggggaa	agaaaaacaaa	gttcctgttt	tcatggaaact	145080
tacattat	tggcgagac	agtaacagac	aaacaaatgt	agcctgtgt	catgtgttac	145140
atgaaaagga	gggttaggggg	ctgggagaga	tttttttttt	tgtagtaggtt	cgaggtgggt	145200
gtcaggaaag	gcctcaactga	ggaggtggca	tttttttttt	acctgagcgc	agcggggggcg	145260
taagcccagg	cagcatgtgg	aggaagagtg	ttcttgggtt	aaggaacaag	gatagaggcc	145320
cgaagctaga	gagctcagca	tgtcaagga	acagcaagcc	ccgtgtggct	gaaatggagt	145380
gagcaaaagga	atgagcagta	gaaggtgagt	gagtttttttt	gtcaccagag	accatggcaa	145440
ggacttgaaa	gtgtcaggga	cacattggaa	gttggagcag	ggaaatgtat	tgatgtttatgt	145500

102/122

tttatgtttt	agtgtttta	aggattgct	ctatcagcta	tttggaaaat	145560
gcttcagaag	gagaaggaga	gaaacaacat	tcttgccata	gtcatagtct	145620
tgatgggtt	tggttaggg	ctggtagtg	aagaccagtc	cagttcggt	145680
ggtagaggca	aaaagattat	atttctacca	gcaagcccat	ctatgaagtt	145740
ttaatttaat	tgagacatgc	ccacataaac	taataaataag	gaatttctgc	145800
aacaccctgt	tatatccgg	ttcttctttt	agttgtccag	atgtctctt	145860
ttttttgggt	gtgttaggagc	ctagagattg	aatttttattca	cccaaaggc	145920
ttactatgtt	ccaggcacta	tgtgtaatgc	caaggatgt	aataagaggg	145980
gtctgtttt	ctccagcttg	gttccctttt	atgaccctt	acttgttaag	146040
atcctacaga	atgtttaaatc	ttctgtactt	tcctggttgt	gttattttagc	146100
ttcccttgaca	tttcttgtaa	actggaagtt	acacctatag	tcttgatgt	146160
catttttagat	tagaacatc	cattgtgtt	atatggtgtt	tttggaaagcc	146220
ttggctctgt	cattaaaatg	ttggctgtat	ggatacacat	aaaatttaac	146280
tttagagatga	gaagaaaagag	gtgcctttt	cttttcaata	taccttttcc	146340
gaactttctt	gccctatgca	tacgttattt	cttaatcatc	cacccatct	146400
ggctttctgt	tgcatttgg	atgaaatcta	gccttcttgc	tgttacctgt	146460
tgctggctc	tatcacatca	cttggaccca	ctcccttcat	ggactgagct	146520
tatcttttat	tcttttgc	aatttttctt	actttgagtg	cctctgcagt	146580
tggctgtggc	aagccctgcc	atgctttca	tgcaaggatg	gttcttctt	146640
tattatctct	tcagagaggg	accttccaa	ctccgatgtat	ctaaaatctt	146700
cactcaactc	cacttcttc	ttttttttt	ctttttatctt	ttttttttt	146760
gagatagggt	tttgcctgt	tgcccaggct	ggaatcacga	ctcactgcag	146820
ttgggctcaa	atgatctct	caccctcagcc	tctcggatgt	ctggaaactgc	146880
caccatactt	ggcttattat	tttacttttt	gttagagacag	gttttcacca	146940
caagctccgt	ccgcaagcaa	tccacatctc	tcagccccc	aaagtattgg	147000
gtgagccact	actcttggc	tattttctt	ttcactgtct	aaaattatct	147060
tttacatact	tgtttatagc	tttatttctt	gctggacatc	gtgcctcaca	147120
caactatttg	ggaggctggg	ttggagaatt	gttggagccc	ggacttcaa	147180
ggcaacaaag	tgagaccctg	tctataaaaa	attgtttaaa	aatttagctgg	147240
acatgcctgt	ggtcccagct	acttgggagg	cagaggtggg	agaatcgctt	147300
ggttgaggcg	acggtgagc	atgattgtgc	cactgcactc	tagcctagtg	147360
accatgtgc	taaaaatcaa	ataaaaatag	tttcttcttc	atgactagaa	147420
atgtgggca	ggagtttgc	tatactattt	ggcactatat	ttcctgattt	147480
cctagcacat	ggtaagact	ccttaaatat	ttattgactg	aattatttaa	147540
tttcattttg	gattatctga	gtgtaagat	tacggattat	atttatgtaa	147600
tttttaaac	ttgggtggcc	tttgcacac	tgacatagac	actaagtttt	147660
ttacttccga	ggataactc	agggccatt	ctttctctaa	tcccccaata	147720
tttagca	tcaagctaat	gcaattctt	gatgatgtat	ctgtgtat	147780
ttctacaaat	gtagaaaattg	aagtctggc	acagtggctc	tcacctgtaa	147840
ttggggaggcc	aaggcgagcg	gatcaactgag	gacaagagtt	aagaccagcc	147900
gttaaaggct	tgcctctatt	aaaaaatacaa	caatttagggc	cgggctgtgt	147960
ataatccca	gcacgttggg	agggcaaggc	aggcagatca	cgaggctagg	148020
ttacctggct	aacacagtga	aacccatct	ctactaaaa	tacaaaaat	148080
ttgtggcacg	cgctttagt	cccagctatc	gggaggctga	ggcaggtgaa	148140
ttggggaggcg	gagggtgcaa	tgagctgaga	ttgcaccgt	gaactccagc	148200
jaggagact	ctgtctctaa	aaaaaaaaaa	aaaaacaatt	agccaggcgt	148260
icgagttacct	gtaatcccag	ctactaggga	ggctgagggga	ggagaatc	148320
jaggtggagt	ttgcagcggg	ctgataatgc	accactacat	tccagcctgg	148380
jagactctgt	ttttttttttt	aaaaaaaagaa	agaaaaagaa	tgaggaatgt	148440
ttctgtgtt	tgttagaaat	cacacagcag	tttagtagca	actacaggc	148500
ataaccacct	tgacaatgtt	tttttttacag	ttcggctccc	tttccctctgc	148560
ttcttattt	agggcagctg	gaaaagattt	tcatcatta	ctagccatata	148620
jagttttgaa	accttgataa	tagacacag	aggaaaagac	tgagttttct	148680
atgtctgtt	ctatggccca	ggctggagtg	cagtgcaccc	atctcagctg	148740
ttgcctccca	ggttcaagca	attttccctc	agcccttcga	gtagctgaga	148800
ttgtcaccac	gcccagctaa	tttttgc	ttgtttcggt	tttgc	148860
ttctgtct	gtcaccagg	ctggatgtca	ttgttgcgt	tttgc	148920
ttctctggg	ttcaagcaat	tcttctgcct	cagccccc	agtagctggg	148980
ttggccacca	tcccttagttc	attttgc	tttttagt	gatgggg	149040
ccaggctgg	tctcgaactc	ctgtatctag	ttgtat	ctgtctc	149100

tcccaaagtg	ctgggattat	tggcacacgc	ctatTTTGT	atTTTtagta	gagacgggggt	149160
ttcaccatgt	tggTTtagact	ggtctcaaAC	ttctgaccTC	aagtGATTTG	cccgccccAG	149220
cctcccaaag	tgctgggatt	acaggcgtga	gccaccgtgc	ccagccaAGA	tttagTTTTG	149280
aaaagagcct	tctgagatTA	tgagaAGGGC	aagcaAGATA	acttaAGAAAG	ttacattAAA	149340
atcatctaag	agacagtgtA	acaAGAAGGA	attgtAAAAT	gatgttatGA	gcacgtGCC	149400
aatgtagtgg	caatCCCTG	tgcttcgata	cattggTGGG	agacAAAACT	gtacttAAAT	149460
tgataaatcc	cttacatgtc	atTTTAAGGA	gcttagactG	actcccata	ttagacATC	149520
agagattct	tttttttttT	tttttttttT	tttttttttT	tttGtgacAG	atTTTgCTC	149580
ttgtggccga	ggctggagtg	caatggcgtg	atctcggtc	accacaACCT	ccacCTCCCA	149640
ggttcaagca	attctcctgc	ctcagccTCC	cgagtagCTG	ggattacAGC	catgcaccAC	149700
cacgcctggc	taattttgtA	tttttagtag	agacggggGT	tctccatGTT	ttggctggTC	149760
tcgaactcct	gacctcaggt	gatcCTCCCG	cctcagccac	ccaaAGTCT	gaaattACAG	149820
gcgtgagCCA	ccgcGCCAG	cccAGAGATT	tctaaACAGA	gttctAACCA	gtgcTTTC	149880
cctgtcgtAG	gaatGAGATA	gaattggagg	ttggagAGAC	tggcatgagg	gacaccAGTC	149940
agccagtggA	attagctgtt	aatgttgata	ggagaAGAAA	aagattcaAA	gttaggtAGT	150000
ggtagcaAGA	attagAGGGA	aggtcggatt	tatgatATGT	ccaaggttGA	attctaAGGT	150060
gaaatttggT	ggcagatttG	atgtgtAAAT	ttggaaAGGT	gattgaggTT	tttaACATG	150120
ggTTTCTAA	catgtcaata	gagtGACTCT	gcaggGGGGC	ctgacGAGAG	aacAGTGCAT	150180
ggggTgATTG	aacAGCCTG	tgaggcTTCA	tgcagAGCAT	ttaacACTGT	gactCTGTAG	150240
actctggTTG	gcagtaAAAT	ttcattAAAC	caatATTTA	acccttaggt	aataataAAA	150300
attgagggAA	aaggatCCAG	gtttgttatt	ttttatGAAT	tcagttattG	aattaaACAG	150360
gaccTTGCT	caagaaATAA	tctaccAAACA	attaACTTGT	tttaaAGCAA	agttAGGAAG	150420
tgagcatgtt	caaatttATA	aataaaaaAG	taagCTGTG	atTCATTCA	tagaaATAGA	150480
gggtggccta	cttcggatGA	ttctcagcat	gtgattACAG	atgtggGCTT	atACATCCTA	150540
gggagttAG	gcgtactCTG	gcttggatAG	atagAGCTC	tttGAAACTC	ttctctcACC	150600
cagctagttt	atataGACTA	gagaACTAGA	atgtAGCAGC	atactCTGTc	ttAGAAGCCC	150660
ttttatataG	gagctggTCT	ggaaGGTTG	aaaACATAAC	aaatgtGTTG	gtgtCTCCCA	150720
atgtattgt	agatttCTAC	ccaaGAGCAT	tatCCTGGTT	agggtttGTT	ttggTTTTGT	150780
tttttttttA	aatgttGCC	acaAAACAAAC	atAGATGTT	agttcttCA	tcaAGTgAGG	150840
agagtagAAAG	aaaAGTCAG	aactCTGAAA	cacCTTTCA	aaagtTTTC	aAGCCATGAT	150900
gtttgcaAGT	taaatGCTCT	gttatgtAA	caatATAATC	atTTTTATT	aatgtAAACAT	150960
tccttagtgt	tttggggTAT	cacacAAAAA	agaATATCCA	tatCTGGAAG	caACAGCTT	151020
taaataAGAG	cattgtggTG	gtggTggTA	tagtggTTT	tttttttttT	tttgagTTG	151080
agtcTcgtc	tgttggccAG	gttggagtG	agtggcaca	tctcagCTCG	tttcaACCTC	151140
tgctcccagg	ttaaGGAAT	tcttcgtcCT	cagCCTCTG	agtagCTGGG	attatAGGCA	151200
cctgtacca	tgcctggCT	atTTTATTAA	tttttagtagA	gacaggTTT	accatGTTG	151260
ccaggctggT	cttgaactCT	taacCTCAGG	tgaatCACCC	acCTCGGCC	cccaaAGTGC	151320
tggattaca	ggcatGAACC	accatGGCCA	gccaAAATAAG	agcattTTTA	atgtAAAATT	151380
atgcatgaaa	tgtacatCA	atTTTGTCT	tgtttactTG	gatccatGTT	ctcacaAGCT	151440
atgaagaaAT	gggtgcAAG	aaataACTGAT	gaggtaAAATC	ctacCTTTAG	gataAAAAGA	151500
tttctgttta	taagtGccAC	cctcatgtAA	gtgaggTTA	aaattttCT	tttctttAGG	151560
tccatgttt	aagcagcatG	gcacatttAT	gttctttac	ccagaatGTA	ccaAGAAAGG	151620
gtgtccctt	cttacacTCT	aacaattGCC	tggtagactG	agtGAAGGT	tcttcAGTCA	151680
gaggcttagA	ccactgaAGG	atatacatGC	attcaAGTT	ccatcAGCCA	gcaggcatCA	151740
gtaatcagtG	tgtagatCAA	aagotCAAAT	gtttcTTCC	ccactggcAG	ttttacttCA	151800
agtagtgGAG	gtttgCTTT	ttaatAGTTA	attaAGTACA	ttgagAGATG	ggaggtGAA	151860
aaaggAAAAT	gttttATTTT	gaccatCTAA	tatGAAAGTA	gttcGgtGTT	aggtatCCAG	151920
tagtgcAC	tggAAAGACAG	ggaATGACAT	gttaATATTG	atAGCAGAG	gttggCCCCAG	151980
gtttttcgT	acatggAAAT	gaaATTCTA	tccaaATAAG	tagaaATTAT	gtgcgtAAAG	152040
cattgttaA	gaggcActGAG	tatgtgcATC	tcgatCCATC	taatGAATAA	cattatCAC	152100
cagttaaat	tatTTTCTT	aggcccAGGA	agagCTAGCT	tggaaAGATTG	ctaaaATGAT	152160
agtcagtGAC	attatGcAGC	aggctcAGTA	tGatcaACCG	ttagAGAAAT	ctacAAAGGT	152220
aaggatgact	tcgttttGtG	taaactAAAAA	agtattATTG	tccAGGTGTA	aaaATAAAA	152280
agaacataAG	gggtttCTT	gcccTTGAAG	gattAACTGC	tgtggggatt	acCTTCTTAT	152340
cataagcaAC	tagaaaATTG	acAAACATAAA	tGAAACAACT	gttgcata	attggacaAT	152400
ggcaataca	ggggAAACCAT	ggAAACCAAA	cagAGCCAG	tagtCTTGT	gaacGAAAGA	152460
gttaaatatC	aaagttcagg	ccaggTgcAG	tggctcAcG	ctgtAAATCCC	agcactTTGG	152520
gaggCCAAGG	cgggtGAATC	acttggAGGT	aggAGTTCAA	gaccAGCTG	gccaACATGG	152580
tgaaaccctG	tcttagcCGG	gtgtggTggc	aggcAccTGT	aatCCCAACT	atTTGGGAGG	152640
ctgaggcagg	agaatcgcTT	gaaccaggGA	ggcggaggTT	gcagtGAGCC	gagatCACAC	152700

cactgcactc	cagcctggc	gacgagcgaa	accccatttc	aaaaaaaaaaa	tcaaagtca	152760
gagagctaa	tttgagtaga	agttgttagga	taaggtagca	aaaaagagga	agtcgcccag	152820
aaagaaagcc	gttagagat	tttagagat	tcccatggat	ccttggctta	ggagtgtatct	152880
gtatatgtt	ggggtaaaa	cgcatgtgc	caggtagaga	accccccaga	aatttaggg	152940
ctgaatgatt	gctggAACAT	agggctaaga	aaagttcatg	gccagaagga	tctggccaga	153000
gttagagac	tttagtaatac	acaaggcatt	gggttagtgc	ttcacagagg	ttatgcctta	153060
ctactgaaga	taaattatgc	ctagagtaca	acgcattca	ccaagttca	aagcaaattt	153120
ttaaagggtc	aaatattacca	acaactgcatt	gccccaaacaa	aggcctaacc	ctctttacag	153180
taacacaaaca	aaatttcagca	cttcacagtg	taaaagtttaga	atgtctgacg	tccaggctgg	153240
gcmcagtggc	tcatgcctgt	aatcccagca	ctttgggagg	ccgaggcagg	tagatgacct	153300
gaggtcagga	gttcaagacc	agcctggcta	acatggtgc	accccgcttca	tattaaaaat	153360
acaaaaactt	agccaggccat	ggtgccggc	acctgtgtatc	ccggctactt	gggaggctga	153420
ggcaggaga	ttgcctgaac	ccagggatg	agggttgc	tgagccgaga	tcgcaccact	153480
gcactctgtt	ctggggaaaa	agagaaaaac	tcaggctcaa	aaaaaaaaaa	aatgtctga	153540
cgtcaatcac	aaatttaccaa	gcatgacatg	agttgaccc	ataaccaggaa	gaaaactcaa	153600
tctatagaaa	cagacccaga	tgtgagaaag	atgatgaatt	tagcagacaa	agaccatcaa	153660
gtggcttattt	taaatattaa	aaatatgttc	agtggccag	gtgcagtggc	tcatgcctgt	153720
aatcccagca	cttggggagg	ccaaggtgg	taggagttca	agaccgctt	ggccaaatatg	153780
gtggaaacccc	ttcttactaa	aaaatcacaaa	aaaattagct	gggcatggtg	gcaggtgcct	153840
atagtcccg	ctatatggga	ggctgaggca	caagaatcac	ttgaacccgg	gaggtggagg	153900
tttagggtgc	agtaagccga	gattgtgcca	cttgtactcc	agcctggaca	acagagttag	153960
actctgtctc	aaaaaaaaaa	aaaaaaaaatg	taaagaaaaac	aagagatataa	tgagaaaaat	154020
gcaaaaatgt	tttaaaaagg	ccaaatggaa	tttctttaaa	aaaaaaaaatc	cagaatggg	154080
ggccggggct	ggtagctcac	gtctataatc	ccagcatctt	gtggggctg	aggcaggcag	154140
atcacctcg	atcggtagtt	caaggccagc	ctgaccaaca	tggagaaacc	tcatctctac	154200
taaaaataca	aaatttagctg	ggcgtgggtgg	cgcattgcct	gtatccctag	ctacttggga	154260
ggctgaggca	ggagaatgtc	ttgaacccgg	gaggcagagg	ttgcggtag	ctgagattgc	154320
accagtgcac	tccagcttgg	gccacaagag	tggaaactccg	tctcaaaaaa	aaaacaaaaaa	154380
aaaacagtag	actcgaagaa	catgtcgatg	ttttctttac	tttaggcagt	aagtgtgacc	154440
ttttcgaggt	gactactta	gttccctcatg	tcctcattag	tagatcagag	aaattcgaca	154500
ccaaaacccc	aaaagaaaaa	ccccctctaa	tcctcattcc	atgattttat	aatgtcatga	154560
agtcctaggg	ctgcgaagga	atactcattc	tctttatcc	gtgttgatac	ctctctgctt	154620
caacccctca	ctgcacattt	gcctataggaa	tgtacttgg	cattcagcat	aaactacctc	154680
acaccattac	tgaattgtt	catgtgcaca	tgtcccatgc	cacaataccg	gggaccttgc	154740
cttccgtgat	atttgtccgc	agtgtgtga	ctacaggagg	gagtcagtga	atgtctgcat	154800
gtgtgtctt	accatccctc	ttgaatatgc	tctagggta	attccttagaa	gtagaattac	154860
tctattgaaa	attggcaata	tttttcatttc	taatatctat	tgccaaatcg	ggaaagcaag	154920
tctggatgcc	agtccctgtt	atatgcctt	tggtaaggat	acgttaacctc	ttaagcttc	154980
tgttcaactca	tatTTTAAACA	agggaaaaat	caatattta	cctcacaaaaa	ttgttagtgc	155040
cttctggctg	tcttaaactc	tggtatata	taaacactaa	gtgttgggt	ccatccctaa	155100
ttttaataaa	tagtgcatt	tttagagaaaa	tgcacccct	cattttcttt	tcttttcttt	155160
tttcagttat	gactcaaaac	tttagataaa	ggaaatctgc	ttgtggaaaa	taagagaact	155220
ttttccctt	ggttggatc	ttcaacacac	ccaatgaaaa	cagcacata	tttctgtatct	155280
gtcaactgtt	tttccaggag	agaatgggg	acaatccctag	acttccacca	taatgcgtt	155340
acctgttaggc	ataatgtat	catatgtgt	tcacacatg	agagtcttaa	agataaaaaaa	155400
tgttattgtt	tacattacta	aaaaattatt	agttttccaa	tggcaataac	ccattttatga	155460
gagtgtttt	gcctactgga	atagacagg	accacatct	ctggggagca	gataagcata	155520
gaactgatac	ttgtatgcaca	ctcgtagtgg	taactcatcc	ctaattcagca	ttgttaagca	155580
ggtgcaggg	gtggttgtt	ttgccttcc	aaagcagggt	agtcaaaaa	accggagggc	155640
aggcagctt	gagtggcagc	gtggtgc	cagcttcagc	ggaacagggt	gagagttat	155700
tatgcagtct	tcttgacagc	ggcattaatt	tggaaaggaaa	ctgacaagtc	atgggtcaag	155760
tttcagttgc	ttccctcc	ctctgtatgc	agttatata	tttcacattt	taattccctcc	155820
tccctgagat	cactataact	aaaaccattc	tctccctgc	taacagaagg	gtgtgaatct	155880
ggtttacttt	gaggcattagg	atttggccct	ttggaaattt	gcactccagg	tacttaactt	155940
tcccttcaga	atacatgtgg	aaagaaaaaa	agaaatgcg	atgactccac	ttttgcccct	156000
gtggcacctt	gaacaaagca	gttcttcca	aattatactt	tttttttttt	taaataaggt	156060
gagcaggatg	actggggaga	gagaaaacatt	tgactttgac	tgcctccccc	attctttgc	156120
gtgagcttgg	aagtgtgcag	ttggcgtct	ttctttctct	ttcttttagga	tagtaagaga	156180
ctcaactca	gcacttctgc	tcagttggct	tctgcatgg	gatcacacag	ccatcagcag	156240
gactgcccag	ttggtgagca	cactccattt	accacgcggc	gccagcgctt	cctcaatgca	156300

105/122

catgatttag	aggaaagaaa	gttctcttag	atgttactgc	ttttgcttag	actttgcaaa	156360
aaaaaaaata	tatatatata	tgtataaata	tataattatt	aatcactttt	gtccttgaga	156420
aagtcttcaa	tgaacagaga	atttattcca	ttgcaatatt	tgattgtata	gagggcacact	156480
gtttcatcg	cagaagaagc	aaaaggctt	tgttaagtt	tttggtaacta	tgtaccacct	156540
ctgttattct	tttaaagctg	aagtattcat	gtacttaaac	catattat	ttaattgtgt	156600
ttgatTTaa	aatatatata	tatgaattct	atttaaaat	gtgtcaactt	tctgctttca	156660
gggcatttat	ggctcttcg	ttgaaatata	ttgatcttt	caaataattt	catttgctt	156720
ctaaaaacc	agaacatgag	ccactactgg	actttgcctt	gtgttgaag	tgtatggcat	156780
aaacccaagg	tttttattag	tcatctatgc	tgtgattaat	tcattttgtt	ctttaacaa	156840
aatatttcca	tccacttcac	attgttcaa	tcttaacag	aaaagaataa	taaaggttat	156900
agaataaaat	gtggtttgg	gcaactcttgc	ctgcctctgc	atgttttgg	ataacaattt	156960
ctcaagact	cttgcgtgtt	taaaactgt	cttcagttt	agataaattt	taatcatttc	157020
tttgtatata	cattttgtgc	ttctgagcta	gagatgcca	gtatgttta	actgcttata	157080
aagagaatag	cagaaaaattt	gagactcgcc	tactttttt	tgccccacct	gctttgagac	157140
acagaagcgg	agtgtggccc	gaaatttata	gccagattta	atatttgatc	taaagttaggt	157200
ccttgcactc	attttaaagt	tggatttga	ttccctccaa	attgagccac	caccatgttc	157260
caggctctgt	gcattgtgc	cacaaaataa	attttccctt	ttggatttt	atgggttcaaa	157320
ataatcgtt	gaacaccctt	catttttatc	atgttggta	cattgacaca	aattgtttaa	157380
aaagaaaaga	tatttagagag	aaagtggatc	cttggtaact	tgatgtgtt	tcatcattcg	157440
gtaagattt	atgaaagtaa	aaagcaatg	tcagccaaat	ccagtgaaaca	gcaataaaaac	157500
agggagtaac	tttttataac	tttttctact	ttgattttca	catttcagtag	agcttttgcg	157560
aatgttaagta	gttttacgt	cttgggggtt	gactagttca	gttagaaattt	ggaggggaag	157620
gtcattctga	attgttaacaa	agtacaaact	tctttctgtt	tttatttaag	tactgagagc	157680
taagcacctg	atgaagtgc	tgacctctct	ccagtgacag	tgtttggta	cctgcctgac	157740
ttcaggagtg	gggtttatgt	ttctacacag	tgacccctt	tctccctc	tcctccctct	157800
tgcccacaca	ccagttgatt	ggacctgggt	tgaactccgt	atccagacag	gccccaaagaca	157860
gtttcttaatg	ttaagaattt	ttggggccggg	cacgggtgtt	catgcctgtt	attgcaacac	157920
tttggggaggc	cgagacaggc	ggatcaactt	aggtcagggg	ttcgaggcc	gcctggccaa	157980
catggtaaaa	ccttgttctt	actaaaaata	caaaaattag	ctgggcatgg	tggcgacacgc	158040
ctgtatcccc	agctacgtgg	gtggctgaga	cagggaaatc	gcttgaacct	ggaggcggag	158100
gttgtcaat	gagccgagac	cgtgtcaact	cattccagcc	ttgggtacag	agggagactc	158160
tgctctccaaa	aataaaaaata	agaaaaaaagaa	ttttgggtca	gttgcagtgg	ctcacgcctg	158220
taattacagc	attttggaaag	gccccaaatg	ggcagatcac	ttgaggacag	gatgtcgaga	158280
ccagccttgg	caacatggtg	aaacctccat	tctactaaaa	agacaaaaatg	tagccagatg	158340
tggtgatggg	cacctataat	cctagcttct	cggggggctg	gggcaggaga	atcacttga	158400
cccaggaagc	agagattgca	gtgagccaag	atcacatctc	tgcactccag	cctgggcaac	158460
agagcaagac	tctgtctcaa	aaaaaaaaga	attttggccag	gcbcgtgtt	tcacgcctgt	158520
aatcccagca	ctttgggggg	ccaaaggcagg	cagatcacga	ggtcaggaga	tcgagattgt	158580
cctggctaac	atggtgaaac	cctgtctctt	ctaaaaatac	aaaacattag	ccgggtgtgg	158640
tggggggcac	ctgtgttccc	agctactagg	gaggctgagg	cagaggaagg	atgtgaaccc	158700
aggaggcggg	gcttgcagta	agccaagatc	gtgccactgc	actacagtct	gggcgcacaga	158760
gtgagactcc	gtctcaaaaa	aaaaaaaat	tttggccggg	tgccgtggca	catgcctgtt	158820
gtcccagcac	tttggggagac	caaagtgggg	ggattaccc	aggtcaggag	tcaagaccca	158880
gtccggccaa	tatggcggaa	ccctgtctct	tactaaaaaa	aataaaaaaa	ttagccaggt	158940
gtgtggccgg	gcacctgggg	aggctgaggc	agggagaaaat	gcttgaaccg	gggaggcaga	159000
ggttgcagta	agccaagatc	gtgccactgc	actccagagc	aagactctt	ctcaaaaaaa	159060
aaaaaaaaag	aattttgcat	ggggaaaggag	agataacttt	caccatctgg	aatgggtctt	159120
ggatgtggca	cttacaaaat	caggagccag	cactgcattt	acaaaacagaa	gcatgtggc	159180
ctgagatagc	aggatccctt	aaaccttgc	agacatccctt	ggtttctgc	tctattccct	159240
cattcttgc	ttggactaca	ttaatctgtc	agttatcttt	ataatgattt	ttgatTTTTT	159300
ttttttgaga	tggagtttcg	cttttgttgc	ccaggcttga	gtgcaatggc	acgatctcgg	159360
ctcaccacaa	cctccaccc	ccagggttca	gtgattctgc	tgcctcagcc	tcctgagtaa	159420
ctgggattac	aggcatgcgc	caccacaccc	ggcttaattt	gtatTTTT	tagagacggg	159480
gtttctccat	gttggtcagg	ctggctctgc	actcccaacc	tcagggtatc	acccctgtctc	159540
ggccctccaa	agtgtcgaaa	ttacaggcg	aagccatgg	acccggctcg	ttttttgatt	159600
ttttgaaacc	agtctgaatg	gagttttttt	aattacgtga	aaggagttt	gctaaaatac	159660
tgccatactg	ccctaattgc	taatgattat	gtattctca	catgcctgtca	agtactgtct	159720
gatttctgga	gaataatttt	tcttttagaa	acttcaactt	agtgcgtatg	tgtatTTCTC	159780
caaaatggta	tcctaactta	atggagctaa	aagacacccc	ttgttttat	aacaaggcgt	159840
tactgaggcc	caggaagggg	agaagtccct	ggcttgtgag	atgatcacca	ttagaactca	159900

106/122

ggcctgggcc	agtgccttt	catgcttc	agatcctcc	aaagaataat	gaagattata	159960
acccttta	gcaattgtaa	taaacccaga	aatagaaaagc	tttttgtta	gagtactgg	160020
agaagttgg	cgggagagat	aattttaca	aaatttgtaa	atacctgcca	attctatata	160080
ctaggcaagg	tctctggcc	tgtaaaaccc	ctcaaggta	caacttgg	ggcccacact	160140
aatagttacc	caactgaggcc	ctctccgggt	gaacatttag	caactagagga	agccccctcg	160200
cttgggcagg	actgggcgtg	gtgcagat	ggagcgggtga	tactgtggat	tctgggcagg	160260
tggagatggc	cagtgtatgc	caataaagga	cacttgagg	agcagtgtga	gtaaaggccc	160320
tgagggcatt	catgttcagg	gaggtgtgc	gcccactggc	ttgcttgc	cacaggagag	160380
tgggtattcc	tgccttagta	actttatgt	aacaagtatt	tcctcagtct	gttcctctca	160440
aactgcctgc	tctggcacat	tcagaatgtc	acagaactca	cctggatgca	ttcagccccct	160500
tgccctaaagg	tgacagtgc	tctcttccc	caccccccaccc	ctcataccac	tgaagcacct	160560
gtcagactgg	cccagtctgt	gggcacaggag	cctagagagg	gcttagtttc	agcttggaaag	160620
gagctgggat	ttaccaagaa	gcaaatgaga	gacggaggatt	gcaacaactg	tgcatttcc	160680
ccagcttcag	ctgactctgt	tatattgact	gtgccttcag	actcatccgt	aagtgacccc	160740
aggctggcct	ctcccacatc	acagtaagaa	ttccacacac	catacaactt	ggaaagaggc	160800
tccagctgaa	ggaagccccca	cacttccc	aagtttttt	tagtcttctc	ttcttggcaa	160860
agagtaccc	ttgtttctt	taattatgt	actattgggt	tagtataat	tcacccattc	160920
agtcaccc	taatgtgcag	gcactgttt	cagggacaca	ggaagaata	aaaacttgca	160980
ggcaccttgg	agcttgatt	ctattgaaga	gtaatggaa	gttggatag	cagctaaact	161040
atgtgttat	tggccaggcg	cagtggctca	cacctgtaat	cccagcacct	tgaggccaa	161100
ggtgggcaga	tcatgaagt	aggagatcga	gaccatctg	gctaacatgg	tgaaaccccg	161160
tctctactaa	aagtaaaaaaa	aaaatttgc	cagggtgtgg	ggcggccg	tgtatgtccca	161220
gctacttggg	aggctgggc	aggagaatgg	tgtgaaccca	ggaggccgaa	attgcagtga	161280
gcccggatgg	caccactgca	ctccagcctg	ggtgacagag	cgagactctg	tctcagaaaa	161340
aaaaaatatg	ctggtagttt	tgattcaaga	tggctttgg	agccatgat	ttaggtctcg	161400
taccaccaa	ggtctactgg	aaaacatcag	gctctctgc	tatagaccca	tagggagac	161460
tgcagccgag	agggggagct	gaagagaagt	gccccttc	tgcctctgtca	gcctcatcct	161520
tccgcaagga	ccagttgtc	tgccacttca	ttcacttgc	gcaagactgg	aggttttcc	161580
tcagggttgt	agcaccttgc	ttacaatgt	tcagcatctt	gatgcctgag	accatcaagg	161640
caagtctctg	aacagggctt	accttagagt	aaggcttaga	agagggcgt	aagtcagtct	161700
cagctccgt	gctctgcaga	gcttggac	atgtgaattc	ttaaaaacaa	gactattgt	161760
cagttaact	atgcatgcag	tataaaat	taaccttgg	aaatcttagc	tagctgttga	161820
gctaattcca	taaagtaatc	agctctgt	tctgcagtg	gtaataataa	tcagcataat	161880
gatgaaacac	tgtgtgtcc	aggcagcgtc	tcatttgatc	cttgcataat	tcttgcataat	161940
actgatttc	tcccttctt	aaacaaagtt	ttttttttt	tttttagagag	ggtctcacta	162000
tgttgccttgc	gctagtcttgc	aattc				162025

<210> 37

<211> 1350

<212> DNA

<213> Homo Sapien

<220>

<221> CDS

<222> (213)...(920)

<300>

<308> GenBank AJ242973

<309> 1999-10-26

<400> 37

gccccccgcgt	cgacgtgaca	gccggtaacgc	ccgggtttgg	gcaacctcga	ttacggccgg	60
cctccaggcc	cgccagcagc	gccccggcc	gccccccccc	gccccctggc	cccccccggtt	120
ccggccgcgg	acccccactt	ctgcgttcc	ggctgcgggt	cegctgcgg	tagcgcgcgtc	180
ccccgggacc	acccttcggc	tggccccc	cc atg ctc tcg gcc acc cgg agg			233

Met Leu Ser Ala Thr Arg Arg

1

5

gct tgc cag ctc ctc ctc cac agc ctc ttt ccc gtc ccg agg atg	
Ala Cys Gln Leu Leu Leu His Ser Leu Phe Pro Val Pro Arg Met	

281

107/122

108/122

ggcaagttga	taaaatgtga	cttatctcctt	aataagttat	ggtgggagtg	gagctgtgcg	1140
gttcctgtg	tcttctgggg	tctgagtgaa	gatagcagg	atgctgttt	cacccttctt	1200
ggtagaagct	aagggtgtgag	ctgggagggtt	gctggacagg	atgggggacc	ccagaagtc	1260
tttatctgtg	ctctctgccc	gccagtgcct	tacaatttgc	aaacgtgtat	agcctcagtg	1320
actcattcgc	tgaaatcctt	cgctttacca				1350

<210> 38

<211> 235

<212> PRT

<213> Homo Sapien

<400> 38

Met	Leu	Ser	Ala	Thr	Arg	Arg	Ala	Cys	Gln	Leu	Leu	Leu	Leu	His	Ser
1									5					15	
Leu	Phe	Pro	Val	Pro	Arg	Met	Gly	Asn	Ser	Ala	Ser	Asn	Ile	Val	Ser
									20					30	
Pro	Gln	Glu	Ala	Leu	Pro	Gly	Arg	Lys	Glu	Gln	Thr	Pro	Val	Ala	Ala
								35					45		
Lys	His	His	Val	Asn	Gly	Asn	Arg	Thr	Val	Glu	Pro	Phe	Pro	Glu	Gly
								50					60		
Thr	Gln	Met	Ala	Val	Phe	Gly	Met	Gly	Cys	Phe	Trp	Gly	Ala	Glu	Arg
								65					80		
Lys	Phe	Trp	Val	Leu	Lys	Gly	Val	Tyr	Ser	Thr	Gln	Val	Gly	Phe	Ala
								85					95		
Gly	Gly	Tyr	Thr	Ser	Asn	Pro	Thr	Tyr	Lys	Glu	Val	Cys	Ser	Glu	Lys
								100					110		
Thr	Gly	His	Ala	Glu	Val	Val	Arg	Val	Val	Tyr	Gln	Pro	Glu	His	Met
								115					125		
Ser	Phe	Glu	Glu	Leu	Leu	Lys	Val	Phe	Trp	Glu	Asn	His	Asp	Pro	Thr
								130					140		
Gln	Gly	Met	Arg	Gln	Gly	Asn	Asp	His	Gly	Thr	Gln	Tyr	Arg	Ser	Ala
								145					155		
Ile	Tyr	Pro	Thr	Ser	Ala	Lys	Gln	Met	Glu	Ala	Ala	Leu	Ser	Ser	Lys
								165					175		
Glu	Asn	Tyr	Gln	Lys	Val	Leu	Ser	Glu	His	Gly	Phe	Gly	Pro	Ile	Thr
								180					190		
Thr	Asp	Ile	Arg	Glu	Gly	Gln	Thr	Phe	Tyr	Tyr	Ala	Glu	Asp	Tyr	His
								195					205		
Gln	Gln	Tyr	Leu	Ser	Lys	Asn	Pro	Asn	Gly	Tyr	Cys	Gly	Leu	Gly	Gly
								210					220		
Thr	Gly	Val	Ser	Cys	Pro	Val	Gly	Ile	Lys	Lys					
								225					235		

<210> 39

<211> 481

<212> DNA

<213> Homo Sapien

<300>

<308> GenBank AW195104

<309> 1999-11-29

<400> 39

ggcattattt	gactgttagt	tttttattaaa	acaaacattt	ctcatagctc	taagcaaagc	60
attagaattc	atcaagcgga	ctcacatctt	ttctctgcac	agagagggc	tgaaaaaggga	120
gagaaagtcc	cttatgtatg	tctagattt	gtaaagcgaa	ggatttcagc	aatgtagtca	180
ctgaggctat	acacgtttgc	aaattgttaag	gcactggcgg	gcagagagca	cagataaagg	240
acttctgggg	tccccccatcc	tgtccagcaa	cctcccaagct	cacaccccttag	cttctaccaa	300
gaagggtgaa	cacagcattcc	ctgttatctt	cactcagacc	ccagaaaaacc	cagggaaaaacc	360
cgacagctcc	actccacca	taacttatta	ggagataagt	cacattttat	caacttgc当地	420

109/122

tcgcgcctcc tatagattat acttcggtaa acccaatctg tataaattcc tttgtacttt 480
g 481

<210> 40
<211> 390
<212> DNA
<213> Homo Sapien

<300>
<308> GenBank AW874187
<309> 2000-05-22

<400> 40
tttttttat tggactgtag gtttttatta aaacaaacat ttctcatagc tctaagcaaa 60
gcattagaat tcatcaagcg gactcacate ttttctctgc acagagaggg ctgaaaagg 120
agagaaagcc ccttatgtat gtctagattt ggtaaagcga aggatttcag cgaatgagtc 180
actgaggcta tacacgttgc caaattgtaa ggcactggcg ggcagagagc acagataaaag 240
gacttttggg ggtccccat tcctgtccag caacctccca gctcacacct tagttctac 300
caagaagggg tgaacacagc atccctgcta tcttcactca gaccggcaga agacacagga 360
aaccgcacag ctccactccc accataactt 390

<210> 41
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 41
agcggtataac aatttcacac agggagctag cttggaagat tgc 43

<210> 42
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 42
gtccaaatata tgcaaacagt tg 22

<210> 43
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 43
agcggtataac aatttcacac agg 23

<210> 44
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

111/122

<223> Oligonucleotide Primer

<400> 50
agccgataac aatbtcacac agggctgaaa tccttcgtt tacc 44

<210> 51
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 51
agccgataac aatbtcacac agg 23

<210> 52
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 52
ctgaaaagg agagaaag 18

<210> 53
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 53
tcccaaagtg ctggaattac 20

<210> 54
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 54
gtccaatata tgcaaacagt tg 22

<210> 55
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Primer

<400> 55
ccccacagcag ttaatccttc 20

112/122

<210> 56
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 56
gcgctcctgt cggtgc
18

<210> 57
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 57
gcctgactgg tggggccc
18

<210> 58
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 58
catgcatttca cggtc
15

<210> 59
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 59
cagagagtac ccctcgaccg tgcatgc
30

<210> 60
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 60
catgcatttca cggtt
15

<210> 61
<211> 30
<212> DNA
<213> Artificial Sequence

113/122

<220>
<223> Oligonucleotide primer

<400> 61
gtacgtacgt gccaaactccc catgagagac 30

<210> 62
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 62
catgcattgc caagg 14

<210> 63
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 63
gcctgactgg tggggcccc 18

<210> 64
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 64
gtgtgcagg tgtaaaacttg taccag 26

<210> 65
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 65
cacggatccg gtagcagccg tagagttg 28

<210> 66
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 66
actggggcatg tggagacag 19

114/122

<210> 67
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 67
gcactttctt gccatgag 18

<210> 68
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 68
tcagtcacga cgtt 14

<210> 69
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 69
cggataacaa tttc 14

<210> 70
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 70
caatttcatc gctggatgca atctgggcta tgagatc 37

<210> 71
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 71
caatttcaca cagcggatgc ttctttggc tctgact 37

<210> 72
<211> 40
<212> DNA
<213> Artificial Sequence

116/122

<210> 78
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 78
tcagtcacga cgttggatgc ccacatgccca cccactac 38

<210> 79
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 79
cggtataacaa tttcggatgc ccgtcaggta ccacg 35

<210> 80
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 80
tcagtcacga cgttggatgc ccacagtggaa gcttcag 37

<210> 81
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 81
gctcataacct tgcaggatga cg 22

<210> 82
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 82
tcagtcacga cgttggatga ccagctgttc gtgttc 36

<210> 83
<211> 34
<212> DNA
<213> Artificial Sequence

117/122

<220>
<223> Oligonucleotide primer

<400> 83
tacatggagt tcggggatgc acacggcgac tctc 34

<210> 84
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 84
tcagtcacga cgttggatgg ggaagagcag agatatacgt 40

<210> 85
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 85
gaggggctga tccaggatgg gtgctccac 29

<210> 86
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 86
tgaagcacctt gaaggatgag ggtgtctgcg 30

<210> 87
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 87
cggataacaa ttccggatgc tgcgtgatga tgaaatcg 38

<210> 88
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 88
gatgaagctc ccaggatgcc agaggc 26

118/122

<210> 89
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 89
gccggccggtg taggatgctg ctgggtgc 27

<210> 90
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide Template

<400> 90
cgcagggttt cctcgtcgca ctgggcattt g 31

<210> 91
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Biotinylated primer

<400> 91
tgcttatccc tgttagctacc ctgtcttggc cttgcagatc caa 43

<210> 92
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 92
agcggataac aatttcacac aggccatcac accgcggtaac tg 42

<210> 93
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 93
cccagtacg acgttgtaaa acgtcttggc cttgcagatc caag 44

<210> 94
<211> 42
<212> DNA
<213> Artificial Sequence

119/122

<220>
<223> Oligonucleotide primer

<400> 94
agccgataac aatttcacac aggcacatcac accgcggtac tg 42

<210> 95
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 95
ctccagctgg gcaggagtgc 20

<210> 96
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide primer

<400> 96
cacttcagtc gctccct 17

<210> 97
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Biotinylated primer

<400> 97
cccagtcacg acgttgtaaa acg 23

<210> 98
<211> 100
<212> DNA
<213> Homo sapien

<400> 98
cctttgagaa agggctctgc ttgagttgt aaaaagaaccg ctgcaacaat ctgggctatg 60
agatcaataa agtcagagcc aaaagaagca gaaaaatgt 100

<210> 99
<211> 100
<212> DNA
<213> Homo sapien

<400> 99
cctttgagaa agggctctgc ttgagttgt aaaaagaaccg ctgcaacaat ctgggctatg 60
agatcagtaa agtcagagcc aaaagaagca gaaaaatgt 100

<210> 100
<211> 100
<212> DNA

120/122

<213> Homo sapien

<400> 100
gaattattt tgggtttcta aaactatggt tcccaataaa agtgactctc agcgagcctc 60
aatgctccca gtgctattca tgggcagctc tctgggctca 100

<210> 101

<211> 100

<212> DNA

<213> Homo sapien

<400> 101
gaattattt tgggtttcta aaactatggt tcccaataaa agtgactctc agcaaggcctc 60
aatgctccca gtgctattca tgggcagctc tctgggctca 100

<210> 102

<211> 84

<212> DNA

<213> Homo sapien

<400> 102
taataggact acttctaattc tgtaagagca gatccctgga caggcgagga atacaggtat 60
tttgtcccttg aagtaacctt tcag 84

<210> 103

<211> 84

<212> DNA

<213> Homo sapien

<400> 103
taataggact acttctaattc tgtaagagca gatccctgga caggcaagga atacaggtat 60
tttgtcccttg aagtaacctt tcag 84

<210> 104

<211> 100

<212> DNA

<213> Homo sapien

<400> 104
ctcaccatgg gcatttgatt gcagagcagc tccgagtccg tccagagctt cctgcagtca 60
atgatcacccg ctgtggcat ccctgaggctc atgtctcgta 100

<210> 105

<211> 100

<212> DNA

<213> Homo sapien

<400> 105
ctcaccatgg gcatttgatt gcagagcagc tccgagtcca tccagagctt cctgcagtca 60
atgatcacccg ctgtggcat ccctgaggctc atgtctcgta 100

<210> 106

<211> 100

<212> DNA

<213> Homo sapien

<400> 106
agcaaggact cctgcaaggg ggacagtggg ggccccatg ccacccacta ccagggcacg 60
tggtaacctga cgggcacgtc cagctggggc cagggtgcg 100

122/122

ggataacattt ggctgtaccc cctgggaaag agcagagata tacgtaccag gtggagcacc	60
caggcctgga tcagccccctc attgtatct gggagccctc	100
<210> 114	
<211> 80	
<212> DNA	
<213> Homo sapien	
<400> 114	
tgaaggactt gaaggagaag gtgtctgcgg gagccgattt catcatcacy cagctttct	60
ttgaggctga cacattcttc	80
<210> 115	
<211> 80	
<212> DNA	
<213> Homo sapien	
<400> 115	
tgaaggactt gaaggagaag gtgtctgcgg gagtcgattt catcatcacy cagctttct	60
ttgaggctga cacattcttc	80
<210> 116	
<211> 80	
<212> DNA	
<213> Homo sapien	
<400> 116	
tccagatgaa gctcccagaa tgccagaggc tgctccccgc gtggccctg caccagcagc	60
tcctacaccg gcggccctg	80
<210> 117	
<211> 80	
<212> DNA	
<213> Homo sapien	
<400> 117	
tccagatgaa gctcccagaa tgccagaggc tgctcccccc gtggccctg caccagcagc	60
tcctacaccg gcggccctg	80
<210> 118	
<211> 48	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Hair pin structure	
<400> 118	
cagagagtac ccctcaaccg tgcatgcattt aaacatgcattt gcacggtt	48